

CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organisms invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of *Aspergillus fumigatus* of the invention.

XX
 CC Sequence 20 BP; 6 A; 1 C; 9 G; 4 T; 0 other;
 Query Match 62.9%; Score 13.2; DB 25; Length 20;
 Best Local Similarity 83.3%; Pred. No. 2.2e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CATCTCACCTTGTCT 19
 |||||
 Db 19 CATACCCATCTTGTCT 2

RESULT 13
 ABV90153/c
 ID ABV90153 standard; DNA; 17 BP.

XX AC ABV90153;
 XX AC
 XX 23-DEC-2002 (first entry)

DE Human POSHL1 scanning oligonucleotide SEQ ID NO 866.

XX Human; POSHL 1; SH3 domain; POSH-like signalling protein 1; oncogene;
 KW Rho GTPase; signal transduction; gene expression; cancer; vaccine;
 KW gene therapy; transgenic; ss.

OS Homo sapiens.

XX

PN EP1239051-A2.

XX 11-SEP-2002.

XX 28-JAN-2002; 2002EP-0001165.

PR 30-JAN-2001; 2001WO-US00663.

PR 30-JAN-2001; 2001WO-US00664.

PR 30-JAN-2001; 2001WO-US00665.

PR 30-JAN-2001; 2001WO-US00666.

PR 30-JAN-2001; 2001WO-US00667.

PR 30-JAN-2001; 2001WO-US00668.

PR 30-JAN-2001; 2001WO-US00669.

PR 23-MAY-2001; 2001WO-US00670.

PR 10-OCT-2001; 2001US-0328205.

XX (AEOM-) AEOMICA INC.

XX Shannon M;

XX WPI; 2002-684061/74.

XX Novel human SH3 domain (POSH)-like signalling protein 1 polypeptide,
 PT POSHL-1, useful for treating disorders associated with decreased
 PT expression or activity of human POSHL1 -

XX Example 2; SEQ ID NO 866; 60pp + Sequence Listing; English.

XX The invention relates to an isolated SH3 domain (POSH)-like signalling
 CC protein 1 (POSHL 1) polypeptide (I), comprising a sequence of 730 amino
 CC acids (S1, ABB83999), a sequence having 65% sequence identity to (S1),
 CC (S1) having 95% deviations, especially conservative substitutions or a
 CC fragment of the sequences comprising at least 8 contiguous amino acids.

CC Human POSHL 1 is a proto-oncogene/oncogene product that functions as an
 CC adaptor protein that interacts with Rho family small GTPases as well as
 CC downstream components of the signal transduction pathway. (I) is useful
 CC for identifying a specific binding partner. (I) and nucleic acids (II)

CC encoding (I) are useful for diagnosing, monitoring disease and treating
 CC caused by altered expression of human POSHL1 including diagnosing and
 CC treating cancer, they useful in the development of vaccines and (II) is
 CC useful in gene therapy. (II) is useful for constructing microarrays which
 CC are useful for measuring and for surveying gene expression and creating
 CC transgenic non-human animals capable of producing the proteins. The
 CC present sequence is that of a scanning oligonucleotide useful in examples
 CC of the invention.

CC Note: The present sequence did not form part of the printed
 CC specification, but is based on sequence information supplied to Derwent
 CC by the European Patent Office.

XX
 SQ Sequence 17 BP; 9 A; 1 C; 6 G; 1 T; 0 other;

Query Match 61.0%; Score 12.8; DB 24; Length 17;
 Best Local Similarity 87.5%; Pred. No. 3.2e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCTCACCTTGTCT 19
 |||||
 Db 17 TTCTCACTCTGTCT 2

RESULT 14
 ABV90154/c
 ID ABV90154 standard; DNA; 17 BP.

XX AC ABV90154;

XX 23-DEC-2002 (first entry)

DE Human POSHL1 scanning oligonucleotide SEQ ID NO 867.

XX Human; POSHL 1; SH3 domain; POSH-like signalling protein 1; oncogene;
 KW Rho GTPase; signal transduction; gene expression; cancer; vaccine;
 KW gene therapy; transgenic; ss.

OS Homo sapiens.

XX

PN EP1239051-A2.

XX 11-SEP-2002.

XX 28-JAN-2002; 2002EP-0001165.

PR 30-JAN-2001; 2001WO-US00663.

PR 30-JAN-2001; 2001WO-US00664.

PR 30-JAN-2001; 2001WO-US00665.

PR 30-JAN-2001; 2001WO-US00666.

PR 30-JAN-2001; 2001WO-US00667.

PR 30-JAN-2001; 2001WO-US00668.

PR 30-JAN-2001; 2001WO-US00669.

PR 23-MAY-2001; 2001WO-US00670.

PR 10-OCT-2001; 2001US-0328205.

XX (AEOM-) AEOMICA INC.

XX Shannon M;

XX WPI; 2002-684061/74.

XX Novel human SH3 domain (POSH)-like signalling protein 1 polypeptide,
 PT POSHL-1, useful for treating disorders associated with decreased
 PT expression or activity of human POSHL1 -

XX Example 2; SEQ ID NO 867; 60pp + Sequence Listing; English.

XX The invention relates to an isolated SH3 domain (POSH)-like signalling
 CC protein 1 (POSHL 1) polypeptide (I), comprising a sequence of 730 amino
 CC acids (S1, ABB83999), a sequence having 65% sequence identity to (S1),
 CC (S1) having 95% deviations, especially conservative substitutions or a
 CC fragment of the sequences comprising at least 8 contiguous amino acids.

CC Human POSHL 1 is a proto-oncogene/oncogene product that functions as an
 CC adaptor protein that interacts with Rho family small GTPases as well as
 CC downstream components of the signal transduction pathway. (I) is useful
 CC for identifying a specific binding partner. (II) and nucleic acids (II)
 CC encoding (I) are useful for diagnosing, monitoring disease and treating
 CC caused by altered expression of human POSHL1 including diagnosing and
 CC treating cancer they useful in the development of vaccines and (II) is
 CC useful in gene therapy. (II) is useful for constructing microarrays which
 CC are useful for measuring and for surveying gene expression and creating
 CC transgenic non-human animals capable of producing the proteins. The
 CC present sequence is that of a scanning oligonucleotide useful in examples
 CC of the invention.
 CC Note: The present sequence did not form part of the printed
 CC specification, but is based on sequence information supplied to Derwent
 CC by the European Patent Office.

XX SQ Sequence 17 BP; 8 A; 2 C; 6 G; 1 T; 0 other;
 Query Match 61.0%; Score 12.8; DB 24; Length 17;
 Best Local Similarity 87.5%; Pred. No. 3.2e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCTCACCTTGTCCT 19
 DB 16 TTCTCACTTGTCCT 1

RESULT 15
 AAA79747/C
 ID AAA79747 standard; DNA; 20 BP.
 XX AC AAA79747;
 XX DT 20-NOV-2000 (first entry)
 XX DE Hepatitis B virus related oligonucleotide probe #10.
 XX KW Hepatitis B virus; HBV; Hepatitis A virus; HAV; probe; detection;
 XX KW mutation; high-density gene chip; SS.
 XX OS Hepatitis B virus.

XX CN1252452-A.
 XX PN 10-MAY-2000.
 XX PD 24-SEP-1999; 99CN-0114460.
 XX PF 24-SEP-1999; 99CN-0114460.
 XX PR (UYDO-) UNIV DONGNAN.
 XX PA Sun X, Lu Z, Wang Y;
 XX PI WPI; 2000-443233/39.
 XX DR High-density gene chip making process -
 XX XX Example 1; Fig 15; 19pp; Chinese.

CC The present invention describes a method which comprises making a high-
 CC density gene chip, specifically for making high-density micro-array of
 CC oligonucleotide probes. An oligonucleotide probe selecting process to
 CC seek preferentially length variable and coverage variable probes is
 CC provided to ensure identical cross melting temperature of probes to the
 CC maximum limit, and this can make the cross control of gene chip
 CC relatively simple and raise the reliability of the gene chip detecting
 CC results. The process proposes a specific probe selection method for
 CC detecting target sequence directly, detecting mutation in both specific
 CC and non-specific sites and a probe overall arrangement scheme. AAA79738
 CC to AAA80201 represent oligonucleotide probe sequences which are used in
 CC examples from the present invention.

SQ Sequence 20 BP; 8 A; 1 C; 9 G; 2 T; 0 other;

Query Match 61.0%; Score 12.8; DB 21; Length 20;
 Best Local Similarity 87.5%; Pred. No. 3.2e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCTCACCTTGTCCT 19
 DB 19 TCCTAACCTTGTCCT 4

Search completed: February 20, 2004, 00:55:57
 Job time : 173 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 00:32:29; Search time 1490 Seconds
(without alignments)
342.546 Million cell updates/sec

Title: US-10-085-108-21_COPY_175_195

Perfect score: 21

Sequence: 1 TCATCCTCACCTTGTCTCA 21

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 8380

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estcl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	57.1	21	28	AZ321746
2	11	52.4	19	28	AZ775540
3	10.4	49.5	20	28	AZ829601
4	10.4	49.5	21	28	AZ580960

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	TITLE	JOURNAL	COMMENT
AZ321746/c	21 bp	DNA	linear	GSS	29-SEP-2000					
IM0042N20F	Mouse 10kb	plasmid	UHGCM	library	Mus musculus genomic					
clone UHGCM0042N20 F	genomic	survey	sequence.							
AZ321746										
AZ321746.1	GT:10374795									
GSS.										
Mus musculus	(house mouse)									
Mus musculus										
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.										
1 (base 1 to 21)										
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weiss, R.										
Mouse whole genome scaffolding with paired end reads from 10kb										
plasmid inserts										
Unpublished										
Contact: Robert B. Weiss										
University of Utah										
Genome Center										
Rm. 308, Biomedical										
Polymers Research Bldg.,										
20 S. 2030 E., SLC, UT										
84112, USA										
Tel: 801 585 5606										
Fax: 801 585 7177										

ALIGNMENTS

5	10	47.6	21	28	AZ313243	AZ313243	1M0029H16
C	6	10	21	28	AZ849030	AZ849030	2M0150117
C	7	9.8	13	12	BG926067	BG926067	HNC23-1-E
C	8	9.8	21	28	AZ654730	AZ654730	1M0529P05
C	9	9.8	21	28	AZ657586	AZ657586	1M053D24
C	10	9.8	21	28	AZ785825	AZ785825	2M0030G05
C	11	9.6	19	9	A1696833	A1696833	wc74e09.x
C	12	9.6	19	28	AZ432757	AZ432757	1M0218L14
C	13	9.6	19	28	AZ654214	AZ654214	1M0528H13
C	14	9.6	19	28	AZ941399	AZ941399	2M0201F07
C	15	9.6	20	28	AZ832404	AZ832404	2M0112J16
C	16	9.6	21	28	AZ820567	AZ820567	2M0092B19
C	17	9.6	21	28	AZ978966	AZ978966	2M0255D04
C	18	9.4	19	28	AZ655467	AZ655467	1M0530017
C	19	9.4	20	28	AZ359918	AZ359918	1M0103A01
C	20	9.4	20	28	AZ436762	AZ436762	1M0224G12
C	21	9.4	20	28	AZ579532	AZ579532	1M0367H10
C	22	9.4	21	28	AZ387199	AZ387199	1M0146P20
C	23	9.2	19	28	AZ458806	AZ458806	1M0263012
C	24	9.2	19	28	AZ816318	AZ816318	2M0085E05
C	25	9.2	21	28	AZ648575	AZ648575	1M0517G07
C	26	9.2	19	9	A1648553	A1648553	tz55e07.x
C	27	9	19	28	AZ330741	AZ330741	1M0056G11
C	28	9	19	28	AZ457990	AZ457990	1M0261E11
C	29	9	20	28	AZ307088	AZ307088	1M0008M23
C	30	9	20	28	AZ784073	AZ784073	2M0026B06
C	31	9	20	29	TA158A03P	TA158A03P	
C	32	9	21	28	AZ435931	AZ435931	1M0223L15
C	33	8.8	19	28	AZ413661	AZ413661	1M0197I07
C	34	8.8	19	28	AZ446372	AZ446372	1M0242P18
C	35	8.8	19	28	AZ447197	AZ447197	1M0244A08
C	36	8.8	20	13	B0583464	B0583464	E011979-0
C	37	8.8	20	28	AZ308291	AZ308291	1M0011E10
C	38	8.8	21	28	AZ394677	AZ394677	1M0158A24
C	39	8.8	21	28	AZ972047	AZ972047	2M0245F16
C	40	8.6	16	9	TA37H07P	TA37H07P	
C	41	8.6	16	9	A1075064	A1075064	ou61g11.x
C	42	8.6	19	9	A1360784	A1360784	gx9g907.x
C	43	8.6	19	28	AZ386656	AZ386656	1M0101K12
C	44	8.6	19	28	AZ487921	AZ487921	1M0229K01
C	45	8.6	19	28	AZ485378	AZ485378	1M0312F16

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0042 row: N column: 20
 Seq primer: CGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 21.
 Location/Qualifiers

FEATURES

source
 1. 21
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0042N20"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWB42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 7 a 0 c 14 g 0 t

Query Match 57.1%; Score 12; DB 28; Length 21;
 Best Local Similarity 75.0%; Pred. No. 5.1e+05;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATCTCCACCTTGCTC 20
 ||||| ||||| ||||| |||||
 Db 21 TCCTCTCTCTCTCTCTCTC 2

RESULT 2
 AZ775540/c 19 bp DNA linear GSS 16-FEB-2001
 LOCUS 2M0008H15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC2M0008H15 F, genomic survey sequence.
 AZ775540
 ACCESSION AZ775540
 VERSION AZ775540.1 GI:12902183
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished

JOURNAL
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606

Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0008 row: H column: 15
 Seq primer: CGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 19.
 Location/Qualifiers

FEATURES

source
 1. 19
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0008H15"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWB42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 7 a 0 c 12 g 0 t

Query Match 52.4%; Score 11; DB 26; Length 19;
 Best Local Similarity 73.7%; Pred. No. 1.2e+06;
 Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATCTCCACCTTGCTCCT 19
 ||||| ||||| ||||| |||||
 Db 19 TCCTCTCTCTCTCTCTCTC 1

RESULT 3
 AZ829601 20 bp DNA linear GSS 20-FEB-2001
 LOCUS 2M0107D03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC2M0107D03 R, genomic survey sequence.
 AZ829601
 ACCESSION AZ829601
 VERSION AZ829601.1 GI:12999605
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished

JOURNAL
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0107 row: D column: 03
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 20.

FEATURES

Location/Qualifiers
 1..20
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0107D03"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 7 a 7 c 0 g 6 t
 ORIGIN

Query Match 49.5%; Score 10.4; DB 28; Length 20;
 Best Local Similarity 70.0%; Pred. No. 2e+06;
 Matches 14; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CATCTCTACCCCTGTCTCA 21
 |||||
 Db 1 CATCATCATCATCATCA 20

RESULT 4
 AZ580960
 LOCUS AZ580960 21 bp DNA linear GSS 13-DEC-2000
 DEFINITION clone UUGC1M0369P04 R, genomic survey sequence.
 ACCESSION AZ580960
 VERSION AZ580960.1 GI:11695495
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 21)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0369 row: P column: 04
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 21.

FEATURES

Location/Qualifiers
 1..21
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0369P04"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 2 a 9 c 2 g 8 t
 ORIGIN

Query Match 49.5%; Score 10.4; DB 28; Length 21;
 Best Local Similarity 91.7%; Pred. No. 2e+06;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CACCTCTGTCT 19
 |||||
 Db 8 CACCTCTGTCT 19

RESULT 5
 AZ313243/c
 LOCUS AZ313243 21 bp DNA linear GSS 29-SEP-2000
 DEFINITION clone UUGC1M0029H16 R, genomic survey sequence.
 ACCESSION AZ313243
 VERSION AZ313243.1 GI:10357979
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 21)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0029 row: H column: 16
Seq primer: CACACAGGAAACGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers

FEATURES

source

1. 21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0029H16"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (G14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 0 c 13 g 6 t

ORIGIN
Query Match 47.6%; Score 10; DB 28; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.8e+06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ATCCTCACCC 12

Db 21 ATCCTCACCC 12

RESULT 6

AZ849030/c
LOCUS 2M0150117F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0150117 F, genomic survey sequence.

ACCESSION AZ849030

VERSION AZ849030.1 GI:13032696

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 21)

REFERENCE

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Ross, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL

COMMENT Unpublished

Contact: Robert B. Weiss
University of Utah Genome Center

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0150 row: I column: 17
Seq primer: CGTTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers

FEATURES

source

1. 21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0150117"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (G14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 6 a 0 c 15 g 0 t

ORIGIN

Query Match 47.6%; Score 10; DB 28; Length 21;
Best Local Similarity 72.2%; Pred. No. 2.8e+06;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TCATCCTCACCCCTGTGCC 18

Db 18 TCCTCCTCTCCTCCTCCCC 1

RESULT 7

BG926067
LOCUS HNC23-1-B8.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
DEFINITION sequence.

ACCESSION BG926067

VERSION BG926067.1 GI:14320590

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 13)

REFERENCE

AUTHORS Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J., Sathe, G., Mui, P., Agarwal, P., Badger, A.M., Lee, J.C., Gower, M. and Lark, M.W.

TITLE Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries

JOURNAL

MEDLINE Osteoarthr. Cartil. 9 (7), 641-653 (2001)

21482651

PUBMED
COMMENT

11597177
Contact: Sanjay Kumar
UW2109
GlaxoSmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay.kumar-1@sk.com
Seq primer: T7

FEATURES
source

1. .13
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/clone_lib="HNC (Human Normal Cartilage)"
/note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
Directional"

BASE COUNT 0 a 9 c 0 g 4 t

Query Match 46.7%; Score 9.8; DB 12; Length 13;
Best Local Similarity 84.6%; Pred. No. 3e+06;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 CTCACCCCTTGTC 18
|||||
Db 1 CTCACCCCTTGTC 13

RESULT 8
AZ654730

LOCUS AZ654730 21 bp DNA linear GSS 14-DEC-2000
DEFINITION M0529P05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0529P05 F, genomic survey sequence.

ACCESSION AZ654730

VERSION AZ654730.1 GI:11791876
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 21)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0529 row: P column: 05
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 21.

FEATURES
source

1. .21
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0529P05"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/clone lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 4 a 13 c 0 g 4 t

Query Match 46.7%; Score 9.8; DB 28; Length 21;
Best Local Similarity 84.6%; Pred. No. 3.4e+06;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CATCCTCACCCCTT 14
|||||
Db 9 CATCCCCCACCCCT 21

RESULT 9
AZ657586

LOCUS

DEFINITION AZ657586 21 bp DNA linear GSS 14-DEC-2000
M0533D24R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0533D24 R, genomic survey sequence.

ACCESSION AZ657586

VERSION AZ657586.1 GI:11794732
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 21)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0533 row: D column: 24
Seq primer: CACACGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.

FEATURES
source

1. .21
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0533D24"
/sex="Male"

```

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGCIM library"
/notes="Vector: PWD42nv, Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      2 a      10 c      2 g      7 t
ORIGIN
Query Match      46.7%; Score 9.8; DB 28; Length 21;
Best Local Similarity 84.6%; Pred. No. 3.4e+06;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 CCTCACCCTTGTCT 17
    |||||
Db 3 CCTCCTCCTTCTC 15

```

```

RESULT 10
AZ785825
LOCUS      21 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0030G05R Mouse 10kb plasmid UUGCIM library Mus musculus genomic
clone UUGC2M0030G05 R, genomic survey sequence.
ACCESSION  AZ785825
VERSION     AZ785825.1 GI:12922972
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0030 row: G column: 05
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers
1. .21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0030G05"

```

```

FEATURES
source
1. .21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0030G05"

```

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGCIM library"
/notes="Vector: PWD42nv, Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      6 a      8 c      3 g      4 t
ORIGIN
Query Match      46.7%; Score 9.8; DB 28; Length 21;
Best Local Similarity 84.6%; Pred. No. 3.4e+06;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TCATCTCACCCT 13
    |||||
Db 6 TCGCCTCACCCT 18

```

```

RESULT 11
AI696833/c
LOCUS      19 bp mRNA linear EST 17-DEC-1999
DEFINITION similar to TR:Q01942 Q01942 EXTENSIN ;contains element TAR1
repetitive element ;, mRNA sequence.
ACCESSION  AI696833
VERSION     AI696833.1 GI:4984733
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 19)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-i@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

```

ACCESSION  AI696833
VERSION     AI696833.1 GI:4984733
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 19)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-i@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

```

Trace considered overall poor quality
Insert Length: 1542 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. .19
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2324392"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Panl"

```

```

FEATURES
source
1. .19
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2324392"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Panl"

```


/note="Organ: pancreas; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

```

BASE COUNT      4 a      1 c      14 g      0 t
ORIGIN
Query Match      45.7%; Score 9.6; DB 9; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      5 CCTCACCCCTTGTCTC 20
      ||||| ||| |||||
Db      19 CCTCCCCCTCTCTC 4

```

```

RESULT 12
AZ432757/c
LOCUS      AZ432757      19 bp      DNA      linear      GSS 03-OCT-2000
DEFINITION      IM0218L14F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0218L14 F, genomic survey sequence.
ACCESSION      AZ432757
VERSION      AZ432757.1 GI:10556770
KEYWORDS      GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0218 row: L column: 14
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1..19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0218L14"
/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated

```

with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

BASE COUNT      6 a      0 c      13 g      0 t
ORIGIN
Query Match      45.7%; Score 9.6; DB 28; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      5 CCTCACCCCTTGTCTC 20
      ||||| ||| |||||
Db      19 CCTCTCCTCTCTCTC 4

```

```

RESULT 13
AZ654214/c
LOCUS      AZ654214      19 bp      DNA      linear      GSS 14-DEC-2000
DEFINITION      IM0528H13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0528H13 F, genomic survey sequence.
ACCESSION      AZ654214
VERSION      AZ654214.1 GI:11791360
KEYWORDS      GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0528 row: H column: 13
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1..19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0528H13"
/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number

```

```

FEATURES
source
1..19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0218L14"
/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated

```

```

FEATURES
source
1..19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0528H13"
/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number

```

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 0 c 14 g 1 t

ORIGIN

Query Match 45.7%; Score 9.6; DB 28; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 CTTCCACCCCTTGCTCTC 20

Db 19 CTTCCACCCCTTGCTCTC 4

RESULT 14

AZ941399

LOCUS

DEFINITION 2M0201F07F Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0201F07 F, genomic survey sequence.

ACCESSION AZ941399

VERSION AZ941399.1

KEYWORDS GI:13803898

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished

COMMENT Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0201 row: F column: 07

Seq primer: CGTGTAAACGACGCGCAGT

Class: plasmid ends

High quality sequence stop: 19.

Location/Qualifiers

1. .19

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0201F07"

/sex="Female"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC2M library"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 7 c 0 g 8 t

ORIGIN

Query Match 45.7%; Score 9.6; DB 28; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CATCTCACCCCTTGTC 17

Db 2 CATCTCACCCCTTGTC 17

RESULT 15

AZ832404/c

LOCUS

DEFINITION 2M0112J16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0112J16 R, genomic survey sequence.

ACCESSION AZ832404

VERSION AZ832404.1

KEYWORDS GI:13002312

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished

COMMENT Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0112 row: J column: 16

Seq primer: CACACAGGAACGACTATGACC

Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers

1. .20

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0112J16"

/sex="Male"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 2 c 7 g 7 t
ORIGIN

Query Match 45.7%; Score 9.6; DB 28; Length 20;
Best Local Similarity 75.0%; Pred. No. 4e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 CTCACCCCTTGTCTCTCA 21
| | | | | | |
Db 20 CACACACTTGTGCACA 5

Search completed: February 20, 2004, 02:39:19
Job time : 1494 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 00:42:09 ; Search time 45.5 Seconds
(without alignments)
203.715 Million cell updates/sec

Title: US-10-085-108-21_COPY_175_195
Perfect score: 21
Sequence: 1 TCATCCTCACCCCTGTCTCTCA 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 352324

Minimum DB seq length: 0
Maximum DB seq length: 21

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.8	75.2	20	3	US-09-433-699-31
2	14	66.7	20	3	US-09-226-012-47
3	13.8	65.7	20	1	US-08-317-432A-26
4	13.4	63.8	21	1	US-08-317-432A-27
5	12.8	61.0	20	3	US-09-422-978-10433
6	12.8	61.0	20	4	US-09-429-323-46
7	12.8	61.0	20	4	US-09-702-327-63
8	12.8	61.0	21	4	US-09-422-978-9636
9	12.6	60.0	20	3	US-09-433-699-40
10	12.6	60.0	20	4	US-09-622-277-2
11	12.6	60.0	21	4	US-09-380-836-88
12	12.4	59.0	18	3	US-09-630-706-59
13	12.4	59.0	19	2	US-08-359-705B-22
14	12.4	59.0	19	2	US-08-286-846A-22
15	12.4	59.0	19	2	US-08-457-860A-22
16	12.4	59.0	19	3	US-08-444-622A-22
17	12.4	59.0	19	3	US-08-942-562-22
18	12.4	59.0	19	3	US-09-156-923-22
19	12.4	59.0	21	4	US-09-380-836-90
20	12.2	58.1	17	4	US-09-371-772B-4357
21	12.2	58.1	18	1	US-08-161-673A-5
22	12.2	58.1	18	2	US-08-481-876-5
23	12.2	58.1	18	3	US-09-185-437-5
24	12.2	58.1	18	3	US-08-652-425-3
25	12.2	58.1	18	3	US-09-632-580A-86
26	12.2	58.1	18	4	US-09-422-978-11706
27	12.2	58.1	20	2	US-08-810-599-47

ALIGNMENTS

RESULT 1

US-09-433-699-31
; Sequence 31, Application US/09433699B
; Patent No. 6165786
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cosset
; TITLE OF INVENTION: ANTISENSE MODULATION OF NUCLEOLIN EXPRESSION
; FILE REFERENCE: RTS-0109
; CURRENT APPLICATION NUMBER: US/09/433.699B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 31
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-433-699-31

Query Match 75.2%; Score 15.8; DB 3; Length 20;
Best Local Similarity .89.5%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	TCATCCTCACCCCTGTCTCT 19
Db	2	TCATCCTCACCCCTGTCTCT 20

RESULT 2

US-09-226-012-47/C
; Sequence 47, Application US/09226012
; Patent No. 6207383
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Splawski, Igor
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; TITLE OF INVENTION: SYNDROME GENE
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/226.012
; CURRENT FILING DATE: 1999-01-06
; EARLIER APPLICATION NUMBER: 09/122.847
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 47
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-226-012-47

Query Match 66.7%; Score 14; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.1e+02; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

QY 5 CCTCACCCCTTGCTCA 18
| | | | | | | | | | | | | | | | | | | | | |
DB 16 CCTCACCCCTTGCTCA 3

RESULT 3
US-08-317-432A-26
; Sequence 26, Application US/08317432A
; Patent No. 5710028
; GENERAL INFORMATION:
; APPLICANT: Nurit Eyal and Nir Navot
; TITLE OF INVENTION: A method of quick screening and
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein
; STREET: 2940 Birchtree lane
; CITY: Silver Spring
; STATE: Maryland
; COUNTRY: United States of America
; ZIP: 20906
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead* Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; SOFTWARE: Word for Windows version 3.11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,432A
; FILING DATE: 4-Oct-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/919,872
; FILING DATE: 27-Jul-92
; APPLICATION NUMBER: 08/084,505
; FILING DATE: 1-Jul-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 128/7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; FILING DATE: 4-Oct-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/919,872
; FILING DATE: 27-Jul-92
; APPLICATION NUMBER: 08/084,505
; FILING DATE: 1-Jul-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 128/7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:

; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-317-432A-26

Query Match 65.7%; Score 13.8; DB 1; Length 20;
Best Local Similarity 98.2%; Pred. No. 1e+03; Indels 2; Gaps 0;
Matches 15; Conservative 0; Mismatches 2;

QY 5 CCTCACCCCTTGCTCA 21
| | | | | | | | | | | | | | | | | | | | | |
DB 2 CCTCACCATTCCTCA 18

RESULT 4
US-08-317-432A-27/c
; Sequence 27, Application US/08317432A
; Patent No. 5710028
; GENERAL INFORMATION:
; APPLICANT: Nurit Eyal and Nir Navot
; TITLE OF INVENTION: A method of quick screening and

; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein
; STREET: 2940 Birchtree lane
; CITY: Silver Spring
; STATE: Maryland
; COUNTRY: United States of America
; ZIP: 20906
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead* Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; SOFTWARE: Word for Windows version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,432A
; FILING DATE: 4-Oct-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/919,872
; FILING DATE: 27-Jul-92
; APPLICATION NUMBER: 08/084,505
; FILING DATE: 1-Jul-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 128/7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-317-432A-27

Query Match 65.7%; Score 13.8; DB 1; Length 21;
Best Local Similarity 88.2%; Pred. No. 1e+03; Indels 2; Gaps 0;
Matches 15; Conservative 0; Mismatches 2;

QY 5 CCTCACCCCTTGCTCA 21
| | | | | | | | | | | | | | | | | | | | | |
DB 20 CCTCACCATTCCTCA 4

RESULT 5
US-09-422-978-10433
; Sequence 10433, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET-020CP1
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 10433
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:

; NAME/KEY: primer bind
; LOCATION: 1..21
; OTHER INFORMATION: downstream amplification primer 99-11851 for SEQ 2568, in complement
US-09-422-978-10433

Query Match 63.8%; Score 13.4; DB 4; Length 21;
Best Local Similarity 93.3%; Pred. No. 1.5e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CTCACCCCTTGCTCTC 20
|||||
Db 7 CTCACCCCTTGCTCTC 21

RESULT 6
US-09-429-323-46
; Sequence 46, Application US/09429323A
; Patent No. 6140126
; Patent No. 6140126 6140123
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF Y-BOX BINDING PROTEIN 1 EXPRESSION
; FILE REFERENCE: RTS-0092
; CURRENT APPLICATION NUMBER: US/09/429,323A
; CURRENT FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 46
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-429-323-46

Query Match 61.0%; Score 12.8; DB 3; Length 20;
Best Local Similarity 87.5%; Pred. No. 2.7e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CATCCTCACCCCTTGTC 17
|||||
Db 2 CCTCCTCACCCCTTTC 17

RESULT 7
US-09-702-327-63
; Sequence 63, Application US/09702327
; Patent No. 6426220
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF CALRETICULIN EXPRESSION
; FILE REFERENCE: RTS-0097
; CURRENT APPLICATION NUMBER: US/09/702,327
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 63
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-702-327-63

Query Match 61.0%; Score 12.8; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 2.7e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 CTCACCCCTTGCTCTCA 21
|||||
Db 1 CTCATCTTGCTCTCA 16

RESULT 8
US-09-422-978-9636/c
; Sequence 9636, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET-020CP1
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 9636
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1..21
; OTHER INFORMATION: downstream amplification primer 99-6261 for SEQ 1771, in complement
US-09-422-978-9636

Query Match 61.0%; Score 12.8; DB 4; Length 21;
Best Local Similarity 87.5%; Pred. No. 2.8e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CCTCACCCCTTGCTCTC 20
|||||
Db 21 CCTCACCCCTTGCTTC 6

RESULT 9
US-09-433-699-40
; Sequence 40, Application US/09433699B
; Patent No. 6165786
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF NUCLEOLIN EXPRESSION
; FILE REFERENCE: RTS-0109
; CURRENT APPLICATION NUMBER: US/09/433,699B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 40
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-433-699-40

Query Match 60.0%; Score 12.6; DB 3; Length 20;
Best Local Similarity 78.9%; Pred. No. 3.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCATCCTCACCCCTTGCTCT 19
|||||
Db 1 TCATCCTCAGCCACGTTCT 19

RESULT 10
US-09-622-277-2
; Sequence 2, Application US/09622277
; Patent No. 6521407
; GENERAL INFORMATION:
; APPLICANT: Warenius, Hilmar Meek

APPLICANT: Seabra, Laurence Anthony
TITLE OF INVENTION: METHODS FOR DETERMINING CHEMOSENSITIVITY OF CANCER CELLS BASED UP
FILE REFERENCE: 1417-188
CURRENT APPLICATION NUMBER: US/09/622,277
CURRENT FILING DATE: 2000-10-25
PRIOR APPLICATION NUMBER: PCT/GB99/00500
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: GB 9903035.5
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: GB 9814545.1
PRIOR FILING DATE: 1998-07-03
PRIOR APPLICATION NUMBER: GB 9812151.0
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: GB 9803447.3
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: GB 9803446.5
PRIOR FILING DATE: 1998-02-18
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR and DNA sequencing primer for exon 2/3 antisense
US-09-622-277-2

Query Match 60.0%; Score 12.6; DB 4; Length 20;
Best Local Similarity 78.9%; Pred. No. 3.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 ATCTCTACCCCTTGCTCTCA 21
||| |||||
DB 1 AGCCACACCTTGCTCTCA 19

RESULT 11
US-09-836-88/c
Sequence 88, Application US/09380836
Patent No. 6551775
GENERAL INFORMATION:
APPLICANT: Lifton, Richard P.
APPLICANT: Chang, Sue S.
APPLICANT: Rossier, Bernard C.
TITLE OF INVENTION: Method to Diagnose and Treat Pathological Conditions
TITLE OF INVENTION: Resulting from Deficient Ion Transport such as
TITLE OF INVENTION: Pseudohypoaldosteronism Type-1
FILE REFERENCE: 44574-5018-US
CURRENT APPLICATION NUMBER: US/09/380,836
CURRENT FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/040,171
PRIOR FILING DATE: 1997-03-11
PRIOR APPLICATION NUMBER: PCT/US98/04601
PRIOR FILING DATE: 1998-03-11
NUMBER OF SEQ ID NOS: 106
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 88
LENGTH: 21
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: G-6 forward
OTHER INFORMATION: PCR primer
US-09-830-836-88

Query Match 60.0%; Score 12.6; DB 4; Length 21;
Best Local Similarity 78.9%; Pred. No. 3.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CATCTCTACCCCTTGCTCTC 20
||| |||||
DB 19 CACACTCAGCCTTGCTCTC 1

RESULT 12
US-09-630-706-59
Sequence 59, Application US/09630706
Patent No. 6277640
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Lex M. Cowser
TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION
FILE REFERENCE: RTS-0053
CURRENT APPLICATION NUMBER: US/09/630,706
CURRENT FILING DATE: 2000-08-01
NUMBER OF SEQ ID NOS: 94
SEQ ID NO 59
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-09-630-706-59

Query Match 59.0%; Score 12.4; DB 3; Length 18;
Best Local Similarity 92.9%; Pred. No. 4.1e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CACCTTGCTCTCA 21
||| |||||
DB 2 CACTCTTGCTCTCA 15

RESULT 13
US-08-359-705B-22/c
Sequence 22, Application US/08359705B
Patent No. 5844092
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
APPLICANT: Urfer, Roman
TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/359,705B
FILING DATE: 20-Dec-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286846
FILING DATE: 08/10/94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/215139
FILING DATE: 03/18/94
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0873P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 base pairs

; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-359-705B-22

Query Match 59.0%; Score 12.4; DB 2; Length 19;
Best Local Similarity 92.9%; Pred. No. 4.1e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CTCACCCCTTGCCCT 19
| | | | | | | | | | | | | | | | | | | | |
Db 19 CTCACCCCTTGCCCT 6

RESULT 14

US-08-286-846A-22/c
; Sequence 22, Application US/08286846A
; Patent No. 5877016

GENERAL INFORMATION:

; APPLICANT: Presta, Leonard G.
; APPLICANT: Shelton, David L.
; APPLICANT: Urfer, Roman
; TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/286,846A
; FILING DATE: 05-Aug-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0873P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

; LENGTH: 19 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear

US-08-286-846A-22

Query Match 59.0%; Score 12.4; DB 2; Length 19;
Best Local Similarity 92.9%; Pred. No. 4.1e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CTCACCCCTTGCCCT 19
| | | | | | | | | | | | | | | | | | | | |
Db 19 CTCACCCCTTGCCCT 6

RESULT 15

US-08-457-880A-22/c
; Sequence 22, Application US/08457880A
; Patent No. 5910574

GENERAL INFORMATION:

; APPLICANT: Leonard G. Presta
; APPLICANT: David L. Shelton
; APPLICANT: Roman Urfer

; TITLE OF INVENTION: HUMAN trk RECEPTORS AND NEUROTROPHIC FACTOR
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/457,880A
; FILING DATE:

CLASSIFICATION:

; PRIOR APPLICATION DATA: US/08/444,622
; APPLICATION NUMBER: 19-May-1995

FILING DATE:

; APPLICATION NUMBER: 08/286846

FILING DATE:

ATTORNEY/AGENT INFORMATION:

; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0873P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

; LENGTH: 19 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear

US-08-457-880A-22

Query Match 59.0%; Score 12.4; DB 2; Length 19;
Best Local Similarity 92.9%; Pred. No. 4.1e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CTCACCCCTTGCCCT 19
| | | | | | | | | | | | | | | | | | | | |
Db 19 CTCACCCCTTGCCCT 6

Search completed: February 20, 2004, 02:40:58
Job time : 47.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 00:32:29 ; Search time 1490 Seconds
(without alignments)
342.546 Million cell updates/sec

Title: US-10-085-108-21_COPY_175_195
Perfect score: 21
Sequence: 1 TCATCCTCACCCCTTGCTCTCA 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 8380

Minimum DB seq length: 0
Maximum DB seq length: 21

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estmu:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_ham:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_ptg:*
- 27: em_gss_vrl:*
- 28: gb_gas1:*
- 29: gb_gas2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	12	57.1	21	28	AZ321746 1M0042N20
C 2	11	52.4	19	28	AZ775540 2M0008H15
C 3	10.4	49.5	20	28	AZ829601 2M0107D03
C 4	10.4	49.5	21	28	AZ580960 1M0369P04

C 5	10	47.6	21	28	AZ313243
C 6	10	47.6	21	28	AZ849030
C 7	9.8	46.7	13	12	BG926067 HNC23-1-E
C 8	9.8	46.7	21	28	AZ654730 1M0529P05
C 9	9.8	46.7	21	28	AZ657586 1M0533D24
C 10	9.8	46.7	21	28	AZ785825 2M0030G05
C 11	9.6	45.7	19	9	AI696833 wc74e09.x
C 12	9.6	45.7	19	28	AZ432757 1M0218L14
C 13	9.6	45.7	19	28	AZ654214 1M0528H13
C 14	9.6	45.7	19	28	AZ941399 2M0201F07
C 15	9.6	45.7	20	28	AZ832404 2M0112J16
C 16	9.6	45.7	21	28	AZ820567 2M0092B19
C 17	9.6	45.7	21	28	AZ978966 2M0255D04
C 18	9.4	44.8	19	28	AZ655467 1M0530U17
C 19	9.4	44.8	20	28	AZ359918 1M0103A01
C 20	9.4	44.8	20	28	AZ436762 1M0224G12
C 21	9.4	44.8	20	28	AZ579532 1M0367H10
C 22	9.4	44.8	21	28	AZ387199 1M0146P20
C 23	9.2	43.8	19	28	AZ458806 1M0263O12
C 24	9.2	43.8	19	28	AZ816318 2M0085E05
C 25	9.2	43.8	21	28	AZ648575 1M0517G07
C 26	9	42.9	19	9	AI648553 tz55e07.x
C 27	9	42.9	19	28	AZ330741 1M0056G11
C 28	9	42.9	19	28	AZ457990 1M0261R11
C 29	9	42.9	20	28	AZ307088 1M0008M23
C 30	9	42.9	20	28	AZ784073 2M0026B06
C 31	9	42.9	20	29	TA158A03P
C 32	9	42.9	21	28	AZ435931 1M0223L15
C 33	8.8	41.9	19	28	AZ413661 1M0197I07
C 34	8.8	41.9	19	28	AZ446372 1M0242F18
C 35	8.8	41.9	19	28	AZ447197 1M0244A08
C 36	8.8	41.9	20	13	BQ583464 E011979-0
C 37	8.8	41.9	20	28	AZ308291 1M0011B10
C 38	8.8	41.9	21	28	AZ394677 1M0158A24
C 39	8.8	41.9	21	28	AZ972047 2M0245F16
C 40	8.8	41.9	21	29	TA37H07P
C 41	8.6	41.0	16	9	AI075064 ou6g511.x
C 42	8.6	41.0	19	9	AI360784 qx98g07.x
C 43	8.6	41.0	19	28	AZ358656 1M0101K12
C 44	8.6	41.0	19	28	AZ438791 1M0229K01
C 45	8.6	41.0	19	28	AZ485378 1M0312F16

ALIGNMENTS

RESULT 1
AZ321746/c
LOCUS
DEFINITION
AZ321746
clone UGUC1M0042N20 F, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ321746 21 bp DNA linear GSS 29-SEP-2000
1M0042N20F Mouse 10kb plasmid UGUC1M library Mus musculus genomic
clone UGUC1M0042N20 F, genomic survey sequence.
AZ321746
GI:10374795
GSS
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.,
and Wright, D. Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0042 row: N column: 20
 Seq primer: CGTTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 21.

FEATURES

source

1. .21
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0042N20"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

7 a 0 C 14 g 0 t

BASE COUNT

ORIGIN

Query Match 57.1%; Score 12; DB 28; Length 21;
 Best Local Similarity 75.0%; Pred. No. 5.1e+05;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATCCTCACCTGTCTCTC 20
 ||||| |||||
 Db 21 TCCTCCTCCTCCTCCTCTC 2

RESULT 2

AZ775540/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ775540 19 bp DNA linear GSS 16-FEB-2001
 2M00008H15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC2M0008H15 F, genomic survey sequence.

AZ775540
 AZ775540.1 GI:12902183
 GSS.

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 19)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606

Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0008 row: H column: 15
 Seq primer: CGTTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 19.

FEATURES

source

1. .19
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0008H15"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

7 a 0 C 12 g 0 t

BASE COUNT

ORIGIN

Query Match 52.4%; Score 11; DB 28; Length 19;
 Best Local Similarity 73.7%; Pred. No. 1.2e+06;
 Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATCCTCACCTGTCTCTC 19
 ||||| |||||
 Db 19 TCCTCCTCCTCCTCCTCTC 1

RESULT 3

AZ2829601

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ2829601 20 bp DNA linear GSS 20-FEB-2001
 2M0107D03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC2M0107D03 R, genomic survey sequence.

AZ2829601
 AZ2829601.1 GI:12999605
 GSS.

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 20)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0107 row: D column: 03
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 20.

FEATURES

Location/Qualifiers
 1..20

/organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0107D03"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

7 a 7 c 0 g 6 t
 BASE COUNT
 ORIGIN

Query Match 49.5%; Score 10.4; DB 28; Length 20;
 Best Local Similarity 70.0%; Pred. No. 2e+06;
 Matches 14; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CATCTCTACCCCTTGCTCA 21
 DB 1 CATCATCATCATCATCA 20

RESULT 4

AZ580960

LOCUS

DEFINITION 1M0369P04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0369P04 R, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Robert B. Weiss

University of Utah

Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0369 row: P column: 04

Seq primer: CACACAGGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 21.

FEATURES

Location/Qualifiers

1..21

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0369P04"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

2 a 9 c 8 t
 BASE COUNT
 ORIGIN

Query Match 49.5%; Score 10.4; DB 28; Length 21;

Best Local Similarity 91.7%; Pred. No. 2e+06;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CACCTCTGCTCT 19

DB 8 CACCTCTGCTCT 19

RESULT 5

AZ313243/c

LOCUS

DEFINITION 1M0029H16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0029H16 R, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Robert B. Weiss

University of Utah

Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0029 row: H column: 16
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 21.
 Location/Qualifiers

FEATURES

source
 1. .21
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U08C1M0029H16"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid U08C1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 0 c 13 g 6 t
 ORIGIN

Query Match 47.6%; Score 10; DB 28; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.8e+06;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ATCTCTACCC 12

Db 21 ATCTCTACCC 12

RESULT 6

AZ849030/6
 LOCUS AZ849030 21 bp DNA linear GSS 21-FEB-2001
 DEFINITION 2M0150117F Mouse 10kb plasmid U08C1M library Mus musculus genomic clone U08C2M0150117 F, genomic survey sequence.

ACCESSION AZ849030

VERSION AZ849030.1 GI:13032696

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 21)

REFERENCE

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0150 row: I column: 17
 Seq primer: CGTTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 21.
 Location/Qualifiers

FEATURES

source
 1. .21
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U08C2M0150117"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid U08C1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 6 a 0 c 15 g 0 t
 ORIGIN

Query Match 47.6%; Score 10; DB 28; Length 21;
 Best Local Similarity 72.2%; Pred. No. 2.8e+06;
 Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATCCTCACCTGTGCC 18

Db 18 TCCTCCTCCTCCTCCTCC 1

RESULT 7

BG926067
 LOCUS BG926067 13 bp mRNA linear EST 06-NOV-2001
 DEFINITION HNC23-1-B8.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA sequence.

ACCESSION BG926067

VERSION BG926067.1 GI:14320590

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 13)

REFERENCE

AUTHORS Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J., Sath, G., Mai, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M. and Lark, M.W.

TITLE Identified and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries

JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)

MEDLINE 21482651

```

PUBMED 11597177
COMMENT Contact: Sanjay Kumar
        UW2109
        GlaxoSmithKline
        709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
        Tel: 610-270-7245
        Fax: 610-270-5598
        Email: sanjay.kumar-1@gsk.com
        Seq primer: T7.

FEATURES             Location/Qualifiers
    source            1..13
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /tissue_type="cartilage"
                        /lab_host="E.coli DH10 B"
                        /clone_lib="HNC (Human Normal Cartilage)"
                        /note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
                        Directional"
    BASE COUNT        0 a 9 c 0 g 4 t
    ORIGIN
        1      46.7%; Score 9.8; DB 12; Length 13;
        2      Best Local Similarity 84.6%; Pred. No. 3e+06;
        3      Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
        4
        5
        6      CTCACCCCTTGTC 18
        7      |||||
        8      1 CTCACCCCTTGTC 13
        9
    QUERY MATCH
        46.7%; Score 9.8; DB 28; Length 21;
        5      Best Local Similarity 84.6%; Pred. No. 3.4e+06;
        6      Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
        7
        8
        9      2 CATCCTCACCCCTT 14
        10     |||||
        11     9 CATCCCCACCCCT 21
        12
    RESULT 8
    AZ654730
    LOCUS             21 bp DNA linear GSS 14-DEC-2000
    DEFINITION        1M0529P05F Mouse 10kb plasmid UUC1M library Mus musculus genomic
    clone UUC1M0529P05 F, genomic survey sequence.
    ACCESSION         AZ654730
    VERSION           AZ654730.1 GI:11791876
    KEYWORDS           GSS.
    SOURCE            Mus musculus (house mouse)
    ORGANISM          Mus musculus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    REFERENCE         1 (bases 1 to 21)
    AUTHORS            Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
        Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
        M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
        and Wright,D., Weiss,R.
    TITLE             Mouse whole genome scaffolding with paired end reads from 10kb
        plasmid inserts
    JOURNAL            Unpublished
    COMMENT            Contact: Robert B. Weiss
        University of Utah Genome Center
        University of Utah
        Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
        84112, USA
        Tel: 801 585 5606
        Fax: 801 585 7177
        Email: ddunn@genetics.utah.edu
        Insert Length: 10000 Std Error: 0.00
        Plate: 0529 row: P column: 05
        Seq primer: CGTTGTAAACGACGCCAGT
        Class: plasmid ends
        High quality sequence stop: 21.
    FEATURES           Location/Qualifiers
        1..21
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UUC1M0529P05"
        /sex="Male"
        /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/clone lib="Mouse 10kb plasmid UUC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
    BASE COUNT        4 a 13 c 0 g 4 t
    ORIGIN
        1      46.7%; Score 9.8; DB 28; Length 21;
        2      Best Local Similarity 84.6%; Pred. No. 3.4e+06;
        3      Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
        4
        5
        6      CATCCTCACCCCTT 14
        7      |||||
        8      9 CATCCCCACCCCT 21
        9
    RESULT 9
    AZ657586
    LOCUS             21 bp DNA linear GSS 14-DEC-2000
    DEFINITION        1M0533D24R Mouse 10kb plasmid UUC1M library Mus musculus genomic
    clone UUC1M0533D24 R, genomic survey sequence.
    ACCESSION         AZ657586
    VERSION           AZ657586.1 GI:11794732
    KEYWORDS           GSS.
    SOURCE            Mus musculus (house mouse)
    ORGANISM          Mus musculus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    REFERENCE         1 (bases 1 to 21)
    AUTHORS            Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
        Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
        M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
        and Wright,D., Weiss,R.
    TITLE             Mouse whole genome scaffolding with paired end reads from 10kb
        plasmid inserts
    JOURNAL            Unpublished
    COMMENT            Contact: Robert B. Weiss
        University of Utah Genome Center
        University of Utah
        Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
        84112, USA
        Tel: 801 585 5606
        Fax: 801 585 7177
        Email: ddunn@genetics.utah.edu
        Insert Length: 10000 Std Error: 0.00
        Plate: 0533 row: D column: 24
        Seq primer: CACACAGGAAACAGCTATGACC
        Class: plasmid ends
        High quality sequence stop: 21.
    FEATURES           Location/Qualifiers
        1..21
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UUC1M0533D24"
        /sex="Male"

```

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 10 c 2 g 7 t
 ORIGIN

Query Match 46.7%; Score 9.8; DB 28; Length 21;
 Best Local Similarity 84.6%; Pred. No. 3.4e+06;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CCTCACCCCTTGTCTC 17
 |||||
 Db 3 CCTCTCCCTTCTC 15

RESULT 10
 AZ785825
 LOCUS
 DEFINITION 2M0030G05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0030G05 R, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

plasmid inserts
 Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0030 row: G column: 05
 Seq primer: CACACAGGACACGATGACC
 Class: plasmid ends
 High quality sequence stop: 21.

FEATURES
 source
 1. .21
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0030G05"

/sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 6 a 8 c 3 g 4 t
 ORIGIN

Query Match 46.7%; Score 9.8; DB 28; Length 21;
 Best Local Similarity 84.6%; Pred. No. 3.4e+06;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCATCTCCACCCCT 13
 |||||
 Db 6 TCGCCCTCACCCCT 18

RESULT 11
 AI696833/c
 LOCUS
 DEFINITION

AI696833 19 bp mRNA linear EST 17-DEC-1999
 wc74e09.x1 NCI_CGAP_Panl Homo sapiens cDNA clone IMAGE:2324392 3, similar to TR:Q01942 Q01942 EXTENSIN ;contains element TARI
 repetitive element ;, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Life Technologies catalog #: 11548-013
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 1542 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES
 source
 1. .19
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2324392"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Panl"

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site1: SalI; Site2: NotI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

BASE COUNT 4 a 1 c 14 g 0 t

ORIGIN

Query Match 45.7%; Score 9.6; DB 9; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CCTCACCTTGTCCCTC 20
|||||
Db 19 COTCCCCCCTCCTC 4

RESULT 12
AZ432757/c
LOCUS
DEFINITION
IM021814F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M021814 F, genomic survey sequence.

ACCESSION
AZ432757
VERSION
AZ432757.1 GI:10556770

KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)

ORGANISM

REFERENCE
1 (bases 1 to 19)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.,
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

UNPUBLISHED

JOURNAL

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0218 row: L column: 14

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 19.

FEATURES

source

1. .19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M021814"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 6 a 0 c 13 g 0 t

ORIGIN

Query Match 45.7%; Score 9.6; DB 28; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CCTCACCTTGTCCCTC 20
|||||
Db 19 COTCCCCCCTCCTC 4

RESULT 13

AZ654214/c

LOCUS

DEFINITION

1M0528H13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0528H13 F, genomic survey sequence.

ACCESSION
AZ654214

VERSION
AZ654214.1 GI:11791360

KEYWORDS
GSS.

SOURCE
Mus musculus (house mouse)

ORGANISM

REFERENCE
1 (bases 1 to 19)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.,
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

UNPUBLISHED

JOURNAL

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0528 row: H column: 13

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 19.

FEATURES

source

1. .19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0528H13"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 0 c 14 g 1 t

ORIGIN

Query Match 45.7%; Score 9.6; DB 28; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 CCTACCCCTTGCTCTC 20
|||||
Db 19 CCTCCCTACCCCTC 4

RESULT 14

AZ941399

LOCUS

DEFINITION 2M0201F07F Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0201F07 F, genomic survey sequence.

ACCESSION AZ941399

VERSION AZ941399.1

KEYWORDS GI:13803898

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished

JOURNAL

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0201 row: F column: 07

Seq primer: CGTGTAAACGACGCCAGT

Class: Plasmid ends

High quality sequence stop: 19.

Location/Qualifiers

FEATURES

source

1..19
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0201F07"
/sex="Female"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC2M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (female). Was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 7 c 8 t

ORIGIN

Query Match 45.7%; Score 9.6; DB 28; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.9e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CATCTCACCTTGCTC 17
|||||
Db 2 CATCTCACCTTGCTC 17

RESULT 15

AZ832404/c

LOCUS

DEFINITION 2M0112J16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0112J16 R, genomic survey sequence.

ACCESSION AZ832404

VERSION AZ832404.1

KEYWORDS GI:13002312

SOURCE GSS.

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished

JOURNAL

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0112 row: J column: 16

Seq primer: CACACGGAACGCTATGACC

Class: plasmid ends

High quality sequence stop: 20.

Location/Qualifiers

FEATURES

source

1..20
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0112J16"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 2 c 7 g 7 t
ORIGIN

Query Match 45.7%; Score 9.6; DB 28; Length 20;
Best Local Similarity 75.0%; Pred. No. 4e+06;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 CTCACCCCTTGTCTCA 21
| | | | | | | | | | | | | | | | | | | | | |
Db 20 CACACCTTGTGCACA 5

Search completed: February 20, 2004, 02:39:19
Job time : 1494 secs

> O <
O||O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-10-085-108-21-inv.res made by tport on Mon 23 Feb 104 9:27:16-PST.

Query sequence being compared:US-10-085-108-21' (1-1041)
Number of sequences searched: 69
Number of scores above cutoff: 69

Results of the initial comparison of US-10-085-108-21' (1-1041) with:

File : 6027924.seq
File : 6475783.seq
File : 6680191.seq
File : US09501104B.seq

100-
-
N -
U -
50-
M -
B -
E -
R -
O -
F 10-
-
S -
E 5-
-
Q -
U -
E -
N -
C -
E -
S 0-
SCORE 0 11 1 21 32 42 53 63 74 84 95
STDEV 11 1 0 1 1 1 1 1 1 1 1

PARAMETERS

Similarity matrix Unitary K-tuple 4
Mismatch penalty 1 Joining penalty 30
Gap penalty 5.00 Window size 500
Gap size penalty 0.33
Cutoff score 1
Randomization group 0

SEARCH STATISTICS

Scores: Mean 28 Median 11 Standard Deviation 29.61
Times: CPU 00:00:00.01 Total Elapsed 00:00:01.00
Number of residues: 61607
Number of sequences searched: 69
Number of scores above cutoff: 69

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Sig.	Frame
1. US-09-066-281B-2	Sequence 20, Application	2940	95	191	2.26
2. US-09-468-433C-2	Sequence 20, Application	2940	95	191	2.26
3. US-09-501-104B-2	Sequence 20, Application	2940	95	191	2.26
4. US-09-468-433C-2	Sequence 21, Application	1041	94	339	2.23
5. US-09-501-104B-2	Sequence 21, Application	1041	94	339	2.23
6. US-09-066-281B-1	Sequence 18, Application	1983	81	203	1.79
7. US-09-468-433C-1	Sequence 18, Application	1983	81	203	1.79
8. US-09-501-104B-1	Sequence 18, Application	1983	81	203	1.79
9. US-09-468-433C-2	Sequence 25, Application	1224	67	338	1.32
10. US-09-501-104B-2	Sequence 25, Application	1224	67	338	1.32
11. US-08-845-528C-1	Sequence 1, Application	4031	60	355	1.08
12. US-09-066-281B-1	Sequence 1, Application	4031	60	355	1.08
13. US-09-468-433C-1	Sequence 1, Application	4031	60	355	1.08
14. US-09-501-104B-1	Sequence 1, Application	4031	60	355	1.08
15. US-08-845-528C-9	Sequence 9, Application	4225	60	355	1.08
16. US-09-066-281B-9	Sequence 9, Application	4225	60	355	1.08
17. US-09-468-433C-9	Sequence 9, Application	4225	60	355	1.08
18. US-09-501-104B-9	Sequence 9, Application	4225	60	355	1.08
19. US-08-845-528C-8	Sequence 8, Application	1691	39	201	0.37
20. US-09-066-281B-8	Sequence 8, Application	1691	39	201	0.37

1. US-10-085-108-21' (1-1041)

US-09-066-281B-2 Sequence 20, Application US/09066281B

Initial Score = 95 Optimized Score = 191 Significance = 2.26
Residue Identity = 33% Matches = 213 Mismatches = 402
Gaps = 14 Conservative Substitutions = 0

380 390 400 410 420 430 440
GAGTGGTTGTCGGGGTCCATATCAGTCAGGGCAATGCCAAAATTAGCTCTATCACTGCGCTTCCCG
450 460 470 480 490 500 510
AAGATCATGGGAAAATAGTCCTTATCTCTTGATG-ACAGTCGTGAGCATCTGCTTGTGACAGGCTC
CATTTGTGAGGAGCGCGAATCAAGTTAGCGGGGGGAGAGTCTTAGACCTGGCCCTCTCAGGGTGAGGG
520 530 540 550 560 570 580
TTTGTGTTGATATTGAGGAGAGAACTGCACCACTCAGCCACCTTTTCATCCAGGGATACCTGGGCA
CCCTGAGGAGAACTGAGGAGACCTCCACCATAGAGAGAAACCCCGCCTGTACTGCGTCCCTGAGA
100 110 120 130 140 150 160
590 600 610 620 630 640 650
GGATTCACTTTCTGGCAAGGCATCCCAAGTAGTGTATCTCTCTTCACTGTGGACTCTCATCAATCG
CTGTAAGTCCCGACAGAGGGAATGCCCCAGAGGAGGAGGAGGTCGCGCCCTTAGGATATAAGG
170 180 190 200 210 220 230
660 670 680 690 700 710 720
GGTCCCAAGAGAGGAGGATGAGCAGGAGTCTAGAGGACTCTGGGAGGAGTCTGCGGAGGACCTCGGGAGA
AAGACACTGAGGAGGG---CTGGGGGAGACGCCACCTCAGAGGGCAGATTCCACAGATTCCACCCCTGC
240 250 260 270 280 290 300
730 740 750 760 770 780 790 800
GATCTTGGGAGGACCTTGGGAGGAGTCTCAGGAGGACTCTGGGAGGAGTCTGCGGAGGACCTCGGAGGAGCAT
TCCTCAAGTATCAGCCCTCGTAGAGTCTCCCGACGTAGCTCAGG---CGGGGTGGCAGCATCTATTTCCTGG
310 320 330 340 350 360 370

810	820	830	840	850	860	870	
ATCCTCATCTCTGGAAACCCACGAATGAGGTAGAGGTAAAGGTGAGGACAAGGTGAGGATGAGGACAA							
380	390	400	410	420	430	440	
GTGAGTGGCCTAGGGAGGCGGAGGCGCTTGTTGCTCAGGGTCCCATGGCAAGTCAGCACGGGGAGCTGCCCTCT							
880	890	900	910	920	930	940	
GGAAAGAGGAGGAGGGGGAATAAAAGTGGAAAGGAAAGGAAAGTGGAGGAGGCATCCTCTCCTCCTCATC							
950	960	970	980	990	1000	1010	
GGTTGGCAGAGGAAAGATTCCACAGCCCTGCTGGGATTAAGACTTGAGGAGTCACATGTGCATCAGACGGAC							
530	540	550	560	570	580	590	
GTGAGGCTACCCGACTGCCCCCATGATGATAGATGCTGGAGAGTGCTGCCACGCCCTACTCCCATGTCTC							
1020	1030	1040					
TGAGGCGTGGAAAGGTTTGGAAAGAGAGGCAT							
600	610	620	630	640	650	660	
TCAGGATGCGCGTGTCTCAGGTTTTCCTTAGGCCACGACAGTGTGGAGGCTCGGCCCTCTCTTGAG							
AAGCCGTGA							
670							

2. US-10-085-108-21' (1-1041)
US-09-468-433C-2 Sequence 20, Application US/09468433C

Initial Score = 95	Optimized Score = 191	Significance = 2.26
Residue Identity = 33%	Matches = 213	Mismatches = 402
Gaps = 14	Conservative Substitutions = 0	

```
380      390      400      410      420      430      440
GAGTGTGTCGGGGTCCATATCAGTCAGGCAGGCAATGCCAAAAATTAGCTCTATGAACCTCATGGCTTTCCCG
TGGGAATCTCAGCGATCGGAGG
X
10      20

450      460      470      480      490      500      510
AAGATCATCGGAAATAGTCCCTTATCTTTGATG-ACAGTCGTGAGATCTCTGCTTTGTGACAGGCTC
CATTTGTAGAGGCGCCGAATCAAGTTAGCGGGGGAAGAGTCTTAGACCTGGCCAGTCTCTAGGTTGAGGG
30      40      50      60      70      80      90

520      530      540      550      560      570      580
TTTGTGTTGATATTTGAGGAGAAGAACTCACCACTCAGCCACCTTTTCATCCAGGGCATACCTGGGCAA
CCCTGAGGAAGAACTCAGGGACCTCCACCATAGAGAGAAGAACCCCGSCCTGTACTGCGTCCCGTGAGA
100     110     120     130     140     150     160

590      600      610      620      630      640      650
GATTCACTTCTGGCAAGCATGCGCAAGTAGCTGTATCTCTCTTCACTGTGTGACTCTCATCCCAATCG
CTGGTAGTCCCAAGCAGGGAATGCCCCAGAGAAGGAGGAGGTGCCGGCCCTCTAGGGAATAAATAGG
170     180     190     200     210     220     230

660      670      680      690      700      710      720
GGTCCACAAGAGGGGATGAGCAGAGACTTAGAGACTCTGGGAGAGACTCTGCGGAGACCCCTGGGGAGA
AAGACACTGAGGAGGG---CTGGGGGAACGCCCCACCTCAGAGGGCGAGATCCCCAGAGATTTCCCAACCTGC
240     250     260     270     280     290     300

730      740      750      760      770      780      790      800
GATCTTGGGAGGACCTGGGGAGGAATCTCAGGAGGACTCTGGGAGAGGTGGCATCCACAGCAGCAGGCAAT
TCCTCAAGTATCAGCCCTGTAGAGTCCCACTCAGCTCAGG---CGGGGTGGCAGCCATCTTATTCTCTGG
310     320     330     340     350     360     370     380
```

ATCCTCATCTTTCTGGAAACCCCGAATAGAGTAGAGGGTTAAGGACAAAGGTTAGGATGAGGATGAGGACAA
 GTGAGTGGCGTAGGGAGGCGGAGGCGCTTGGTCTGAGGGTCCCATGCGAAGTCAGCACGCGGAGCTGCCTCT
 380 390 400 410 420 430 440
 GGAAGAGGAGGAGGGAATAAAGTGGAAAGAGAGAGGAGTGGAGGAGCATCTCTCTCTCTCTCATC
 GGTTCGACAGAGGAAGATTCCACAGGCCCTCTGCTGGGATAAGACTGAGGAGTCAATGTGCATCAGAACGACG
 450 460 470 480 490 500 510 520
 950 960 970 980 990 1000 1010
 TATGGAATCCTGTGCACTTACCAAGTCCTCTGTCACTCGGTTCTGGAA-----GTCTTCTCTCAAGAC
 GTGAGGCTACCCGACTGCCCCCATGGTAGAGTCTGGAGGTGGCTGCCACGCCCTACCTCCCACTGTCTC
 530 540 550 560 570 580 590
 TGAGCGGTGGAAGGTTTGGAAAGAGGCAT
 1020 1030 1040
 TCAGGATGTGCGGTTGTCTCTGAGGTTTCTTAGGCCAGCAGAGTGGTGGAGGCTGGCCCTCTCTGAG
 600 610 620 X 630 640 650 660
 AAGCCGTGA
 670

3. US-10-085-108-21' (1-1041)
 US-09-501-104B-2 Sequence 20, Application US/09501104B

Initial Score = 95 Optimized Score = 191 Significance = 2.26
 Residue Identity = 33% Matches = 213 Mismatches = 402
 Gaps = 14 Conservative Substitutions = 0

380 390 400 410 420 430 440
 GAGTGTCTCGGGTCCATACAGTCAGGCGCAATGCCAAAATTAGCTCTATGAATCATGGCTTTCCCG
 TGGGAATCTGACCGATCGGAGG
 X 10 20
 450 460 470 480 490 500 510
 AAGATCATGGGAAATAGTCCATTACTTTTGATG-ACAGTCGTGAGCATCTGCGCTTTGAGCAGGCTC
 CATTTGTAGGAGGCGCGAATCAAGTTAGCGGGGAGAGTCTTAGACCTGGCCAGTCTCAGGCGTAGGG
 520 530 540 550 560 570 580
 TTTTGTGTATTTGAGGAGAGAAATGACCACTCAGCCACTTTTTCATCAGGAGTACCTTGGGAA
 CCTGAGGAAGACTGAGGGAACCTCCCACTAGAGAGAAGAAACCCGCGCTGTACTGCGTGCCTGAGA
 100 110 120 130 140 150 160
 590 600 610 620 630 640 650
 GGATTCACCTTTCTGCAAGGCATGCCAGTACTGTATCTCTTCACTCTGACTCCTCATCTCATGTCG
 CTGTTAGTCCACAGACGGAATAATGCCCCAGAGAGGAGGAGTGTCGCGCTCTTAGGGAATAATAGG
 170 180 190 200 210 220 230
 660 670 680 690 700 710 720
 GGTCACAAAGACGGGATGACGAGGCTAGAGGACTCTGGGAGGACTCTGGGAGGACCTCTGGGAGA
 AAGACACTGAGAGG---CTGGGGAGACGCCCACTCAGAGGCGAGATTCCAGAGATTCCCACTCCGTCG
 240 250 260 270 280 290 300
 730 740 750 760 770 780 790 800
 GATCTTGGGAGACCTCGGAGGAATCTCAGAGGACTCTGGGAGAGGTGGCATCCAGCAGCAGGCAT
 TCCTCAAGTATCAGCCCTGTAGACTCCCACTCAGTCTCAG---CGGGTGGAGCCATCTTATTCCTGG
 310 320 330 340 350 360 370
 ATCCTCATCTTCTGGAAACCCCGAATAGAGTAGAGGTTAAGGACAAAGGTTAGGATGAGGACAA


```
430      440      450      460      470      480      490
ACTCATGGGCTTTCCCGAAGATCATGGGAAATAGTCTTATCTTCTTGATGACAGTGGTCAAGCATCTCTG
      10      20
TGGGAATCTGACGGATCGGAGG
X
500      510      520      530      540      550      560
CCTTTGTACAGGCTCTTTTGTGATATTTGAGGAGAAACTGCAACTCAGCCACC-----TTTTC
      10      20
CAITTTGTAGGAGGCGCGAAATCAAGTTAGCGGGGGAAGAGTCTTAGACCTGGCCAGTCTCAGGGTGAGG
      30      40      50      60      70      80      90
570      580      590      600      610      620      630
ATCCAGGGCAATCTGGCAAGGATCTACTTTCTGGCAAGGATGCGCAAGTAGTGTATCTCTCTTCACT
      10      20
CCCTTGAGGAGAACTGAGGAGCTCCCAACCATAGAGAGAAACCCCGGCTGTACTGCGCTGCCGTGAGA
      30      40      50      60      70      80      90
CAITTTGTAGGAGGCGCGAAATCAAGTTAGCGGGGGAAGAGTCTTAGACCTGGCCAGTCTCAGGGTGAGG
      100      110      120      130      140      150      160
640      650      660      670      680      690      700
GCTGGACTCTCTATCCCAATCGGGTCCCAAAAGAGGGGATGAGCAGGAGTCTAGAGGACTCTGGGAGGACT
      10      20
CTGTGCTCTCCAGGAAACCAAGGTGGTGAACCACTGGTGTGAGGACACAGCCCTAAAGTCAGCACAGCAGGA
      30      40      50      60      70      80      90
170      180      190      200      210      220      230
CTGGAGGAGGACCCCTGGGAGAGATCTTTGGGAGGACCCCTGGGAGGAATCTCAGGAGGACTCTGGGAGAGG
      10      20
GGCCAGGAGGAGTGGCAGGAGTCAAGGCTCTGGATCTCATCTCATCTCATCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
240      250      260      270      280      290      300      310
TGGCATCCAGCAGCAGCATATCTCTATCTCTTGGAAACACCCAGAGATGAGAGTAGAGGTAAAGGTGAGGA
      10      20
TGCTCTCTGAGAAAGTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
790      800      810      820      830      840      850
CAAGGTTGAGGATGAGGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
      10      20
CTCAGTTGAGTTAGAACTGGGTAGATGCACAGCATCCACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG
      30      40      50      60      70      80      90
930      940      950      960      970      980      990
GGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      10      20
CTCTTCCACTTTGTACTTAGTATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
1000      1010      1020      1030      1040
GTCTTCTCTAAAGCTGAGGCGTGAAGTTTGGAAAGAGAGGCAAT
      10      20
GGTCTCTGAGGAGGAGGAGTGGCCCTCTGGTGTGATACCAATCTTACCGAGAGCATCCAGTAGTCTCTCA
      30      40      50      60      70      80      90
CAGGGTCTCTCCACAGGGTCTCTTC
      100      110      120      130      140      150      160
```

8. US-10-085-108-21', (1-1041)
US-09-501-104B-1 Sequence 18, Application US/09501104B

Initial Score = 81 Optimized Score = 203 Significance = 1.79
Residue Identity = 37% Matches = 215 Mismatches = 346
Gaps = 10 Conservative Substitutions = 0

```
430      440      450      460      470      480      490
ACTCATGGGCTTTCCCGAAGATCATGGGAAATAGTCTTATCTTCTTGATGACAGTGGTCAAGCATCTCTG
      10      20
TGGGAATCTGACGGATCGGAGG
X
500      510      520      530      540      550      560
CCTTTGTACAGGCTCTTTTGTGATATTTGAGGAGAAACTGCAACTCAGCCACC-----TTTTC
      10      20
CAITTTGTAGGAGGCGCGAAATCAAGTTAGCGGGGGAAGAGTCTTAGACCTGGCCAGTCTCAGGGTGAGG
      30      40      50      60      70      80      90
570      580      590      600      610      620      630
ATCCAGGGCAATCTGGCAAGGATCTACTTTCTGGCAAGGATGCGCAAGTAGTGTATCTCTCTTCACT
      10      20
CCCTTGAGGAGAACTGAGGAGCTCCCAACCATAGAGAGAAACCCCGGCTGTACTGCGCTGCCGTGAGA
      30      40      50      60      70      80      90
CAITTTGTAGGAGGCGCGAAATCAAGTTAGCGGGGGAAGAGTCTTAGACCTGGCCAGTCTCAGGGTGAGG
      100      110      120      130      140      150      160
640      650      660      670      680      690      700
GCTGGACTCTCTATCCCAATCGGGTCCCAAAAGAGGGGATGAGCAGGAGTCTAGAGGACTCTGGGAGGACT
      10      20
CTGTGCTCTCCAGGAAACCAAGGTGGTGAACCACTGGTGTGAGGACACAGCCCTAAAGTCAGCACAGCAGGA
      30      40      50      60      70      80      90
170      180      190      200      210      220      230
CTGGAGGAGGACCCCTGGGAGAGATCTTTGGGAGGACCCCTGGGAGGAATCTCAGGAGGACTCTGGGAGAGG
      10      20
GGCCAGGAGGAGTGGCAGGAGTCAAGGCTCTGGATCTCATCTCATCTCATCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
240      250      260      270      280      290      300      310
TGGCATCCAGCAGCAGCATATCTCTATCTCTTGGAAACACCCAGAGATGAGAGTAGAGGTAAAGGTGAGGA
      10      20
TGCTCTCTGAGAAAGTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
790      800      810      820      830      840      850
CAAGGTTGAGGATGAGGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
      10      20
CTCAGTTGAGTTAGAACTGGGTAGATGCACAGCATCCACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG
      30      40      50      60      70      80      90
930      940      950      960      970      980      990
GGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      10      20
CTCTTCCACTTTGTACTTAGTATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
1000      1010      1020      1030      1040
GTCTTCTCTAAAGCTGAGGCGTGAAGTTTGGAAAGAGAGGCAAT
      10      20
GGTCTCTGAGGAGGAGGAGTGGCCCTCTGGTGTGATACCAATCTTACCGAGAGCATCCAGTAGTCTCTCA
      30      40      50      60      70      80      90
CAGGGTCTCTCCACAGGGTCTCTTC
      100      110      120      130      140      150      160
```

```
CTTTGTGACAGGCTCTTTTGTGATATTTGAGGAGAAACTGCAACTCAGCCACC-----TTTTC
      10      20
CAITTTGTAGGAGGCGCGAAATCAAGTTAGCGGGGGAAGAGTCTTAGACCTGGCCAGTCTCAGGGTGAGG
      30      40      50      60      70      80      90
570      580      590      600      610      620      630
ATCCAGGGCAATCTGGCAAGGATCTACTTTCTGGCAAGGATGCGCAAGTAGTGTATCTCTCTTCACT
      10      20
CCCTTGAGGAGAACTGAGGAGCTCCCAACCATAGAGAGAAACCCCGGCTGTACTGCGCTGCCGTGAGA
      30      40      50      60      70      80      90
CAITTTGTAGGAGGCGCGAAATCAAGTTAGCGGGGGAAGAGTCTTAGACCTGGCCAGTCTCAGGGTGAGG
      100      110      120      130      140      150      160
640      650      660      670      680      690      700
GCTGGACTCTCTATCCCAATCGGGTCCCAAAAGAGGGGATGAGCAGGAGTCTAGAGGACTCTGGGAGGACT
      10      20
CTGTGCTCTCCAGGAAACCAAGGTGGTGAACCACTGGTGTGAGGACACAGCCCTAAAGTCAGCACAGCAGGA
      30      40      50      60      70      80      90
170      180      190      200      210      220      230
CTGGAGGAGGACCCCTGGGAGAGATCTTTGGGAGGACCCCTGGGAGGAATCTCAGGAGGACTCTGGGAGAGG
      10      20
GGCCAGGAGGAGTGGCAGGAGTCAAGGCTCTGGATCTCATCTCATCTCATCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
240      250      260      270      280      290      300      310
TGGCATCCAGCAGCAGCATATCTCTATCTCTTGGAAACACCCAGAGATGAGAGTAGAGGTAAAGGTGAGGA
      10      20
TGCTCTCTGAGAAAGTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
790      800      810      820      830      840      850
CAAGGTTGAGGATGAGGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
      10      20
CTCAGTTGAGTTAGAACTGGGTAGATGCACAGCATCCACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG
      30      40      50      60      70      80      90
930      940      950      960      970      980      990
GGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      10      20
CTCTTCCACTTTGTACTTAGTATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
1000      1010      1020      1030      1040
GTCTTCTCTAAAGCTGAGGCGTGAAGTTTGGAAAGAGAGGCAAT
      10      20
GGTCTCTGAGGAGGAGGAGTGGCCCTCTGGTGTGATACCAATCTTACCGAGAGCATCCAGTAGTCTCTCA
      30      40      50      60      70      80      90
CAGGGTCTCTCCACAGGGTCTCTTC
      100      110      120      130      140      150      160
```

9. US-10-085-108-21', (1-1041)
US-09-468-433C-2 Sequence 25, Application US/09468433C

Initial Score = 67 Optimized Score = 338 Significance = 1.32
Residue Identity = 36% Matches = 393 Mismatches = 628
Gaps = 42 Conservative Substitutions = 0

```
X
      10      20      30      40      50      60
CTAGGATGATCTGATAGCT-----TGGATAAACTCTTAAGACTTCTCTTGGCTCTCTCTCTCTCTCTCT
      10      20      30      40      50      60
ATCCCTCGGGGTCAAGAGTGAAGTCTCCGTACCTGTGAGAAAGCCCAAGAGACCAATGGTTCAGCCACAGG
      30      40      50      60      70      80
70      80      90      100      110      120      130
GGGCTCTTGAACCCCAAAATCTAAGCTGAGGAGGAGTGTGGGCACTCCCGGCTCTCTCTCTCTCTCTCT
      10      20      30      40      50      60      70
CTCAGCGGTCTCCAGGCCCACTGCAGAGAAAGCAGGAGAGTCCCACTCTCTCTCTCTCTCTCTCTCTCTCT
      30      40      50      60      70      80      90
140      150      160      170      180      190      200
TTCTCTGCAACCAATGATAGTAGCAGGCTTCTGGATCCCAATATATAAGTCTCTCTCTCTCTCTCTCTCT
      10      20      30      40      50      60      70
```

[illegible]

[illegible][illegible]

US-10-066-281B-1 Sequence 1, Application US/09066281B (I-1041)

AGCACTTTGCTTTTGGGGAGCCCGAGGAGCTCTCTCACTAAAGTTTGGGTGAGGAACATTACCTAGAGTAC
 3150 3160 3170 3180 3190 3200 3210
 CATCCGAGCAGGAGCATATCCCTCATCTTCTCTGAAACAC-----CCAGAATGAGAGTAGAGGTTAAGGG
 790 800 810 820 830 840
 GGAGGTTGCCCAACTCTTCTCTCTCTCGTTTCCGAATCTCTGTGGGGTCCAGAGAGCTCATTCAGAGTCATTA
 3220 3230 3240 3250 3260 3270 3280
 850 860 870 880
 TGAAGCAAGGGTTCAGGATGAGGCAAGAGGAGGAGGG-----CA
 AGAGGAAGTGTAGTAGGTTTGGCCATCTCTAAGATACCGTCCCTATTACCTTTCCATCTCTTACAGG
 3290 3300 3310 3320 3330 3340 3350
 ATGCTTTGAAGATGTGGAGAGAGAGCCGAGGCCATAATTGACACCAAGATGATTCGACTGCCACAGAAA
 3360 3370 3380 3390 3400 3410 3420 3430
 900 910 920 930 940 950 960
 ATAAAAGTGGAAAGAGAGAGGAAGTGGAGGAGCATCTCTCTCTCATCTATGAATCTCTGTGCAT
 CTACCAAGTCTCTGTACACTCGGGTCTTGAAGTCTTCTCTCAAGCTGAGCGTGAAGGTTTGGAAAGA
 GTGCAAGCTCCAGTGTATGTCCCCAGCTTCTCTTCTGAGTGAAGTCTAAGGCGAGATCTTCCCTCTGAGT
 3440 3450 3460 3470 3480 3490 3500
 X
 GAGGCAAT
 TTGAAGGGGCGAGTCGAGTTTCTACGTGTGTGGAGGCGCTGGTTGAGGCTGGAGAGAA
 3510 3520 3530 3540 3550 3560
 13. US-10-085-108-21' (1-1041)
 US-09-468-433C-1 Sequence 1, Application US/09468433C
 Initial Score = 60 Optimized Score = 355 Significance = 1.08
 Residue Identity = 37% Matches = 419 Mismatches = 619
 Gaps = 71 Conservative Substitutions = 0
 X 10 20
 CTAGGGATGATCTGGATAGCT
 CTGCTCAGTCTCTCTCCAGAGACCTGTGAGCTCTTCTCTCTACACTTTAGCGAGTCTTCTCCAAAGTT
 2360 2370 2380 2390 2400 2410 2420
 30 40 50 60 70 80 90
 TGGATAAAACTCTAAGACTTCTTGTCTGG--CCTCTGAATGGGCTCTTGGAGCCACAAAATTCATACAC
 CCATGAGAGTCTCTCAGTCTCTCTGAGGGGCTGCCAGTCTCTCTCCAGAGTCTCTGTAGCTCTCTTCCC
 2430 2440 2450 2460 2470 2480 2490
 100 110 120 130 140 150 160
 GTGGAGGAGCACTGTGGGCACTTCCC--GGTACTCAGGTACTTCTCTGACCCCAATGATNTAGTCAGGACAG
 CCTCTCCACTTCATCAGGTCCTTCCAGAGTCTCTCTGTAGCTCTCTCCCTCTCCACTTCATCAGGTC
 2500 2510 2520 2530 2540 2550 2560 2570
 170 180 190 200 210 220 230
 CTTTCTGGGATCCCATATATAAGTGCTCCCTCCAGACACACCCTATTGCATCAACACTTCCAGAT
 TTTCGAAGAGTCCCTTGAGAGTCTCTCCAGAGTCTCTGTGATCTCTCTCTCTCCCTCACTTCATTGAGCC
 2580 2590 2600 2610 2620 2630 2640
 240 250 260 270 280 290
 GACCTCTCGGGACACAGCTGGCCCTTTATGAAGATCATACTGATAAGA-----ATCAGGAGACAGCT
 CATTCAGTGAAGATCCAGAGCCAGTAGATGAATATACAAGTTCTCTCGACACCTTGTAGAGATGATT
 2650 2660 2670 2680 2690 2700 2710
 300 310 320 330 340 350 360 370

	670	680	690	700	710	720	730
--	-----	-----	-----	-----	-----	-----	-----


```
820 830 840 850 860 870 880
GGGAAGTGTGAGTGAATAGGGGTGTGCTGGAGGAGGACCTTTATATATGGGATCCCAAGAGCTGC
1080 1090 1100 1110 1120 1130 1140
GGGAAGTGTGAGTGAATAGGGGTGTGCTGGAGGAGGACCTTTATATATGGGATCCCAAGAGCTGC
1080 1090 1100 1110 1120 1130 1140
TCACATACATATGGGTGCAGAGAAGTACCTGGAGTACCGGGAGTGCCCAACAGTGTCTCCCAAGTATG
1150 1160 1170 1180 1190 1200 1210 1220
TCACATACATATGGGTGCAGAGAAGTACCTGGAGTACCGGGAGTGCCCAACAGTGTCTCCCAAGTATG
1150 1160 1170 1180 1190 1200 1210 1220
AAATTTTGTGGGTGCAGAGCCCAATTCAGAGGCCAGCAAGA--GAAGTCTTAGAGTTTTTATCCAGCTAT
1230 1240 1250 1260 1270 1280 1290
AAATTTTGTGGGTGCAGAGCCCAATTCAGAGGCCAGCAAGAAGTCTAGAGTTTTTATCCAGCTATG
1230 1240 1250 1260 1270 1280 1290
1030 X
CCAGTATCATCCCTAG
1030
ACAACTATGTTCTAGTTCCTTTCCATCCTGGTACAGAGTGTCTTGAAGATGTGGAAGAGAGAG
1300 1310 1320 1330 1340 1350
ACAACTATGTTCTAGTTCCTTTCCATCCTGGTACAGAGTGTCTTGAAGATGTGGAAGAGAGAG
1300 1310 1320 1330 1340 1350
30 40 50 60 70 80 90
CACCCCTCAGCTTTGAGGAAGACTTCCAGAACCCGAGTGTGACAGAGACTTGTAGATGACAGAGTCCA
1310 1320 1330 1340 1350 1360 1370 1380
CATTCGCAAGCTTTGACAAACACTCCCGGACTCAGTTGAGTTAGAGACTGGGTAGANGACAGCATCCA
1310 1320 1330 1340 1350 1360 1370 1380
100 110 120 130 140 150 160
TAGATGAGGAGGAGGAGTCCCTCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1390 1400 1410 1420 1430 1440
CAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1390 1400 1410 1420 1430 1440
170 180 190 200 210 220 230
CCTTGTCTCTATCTCTACCCCTGTGCTTACCCCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1450 1460 1470 1480 1490
TCT-----CCACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1450 1460 1470 1480 1490
240 250 260 270 280 290 300 310
ATATGCTGCTGCTGGAGTGCACCTCTTCCCAAGAGTCTCTAGACTCTCTCTCTCTCTCTCTCTCTCT
1500 1510 1520 1530 1540
AGGTGCCCTCTGCTGATACCAATCTTACCGAGA-----GCATTCCCAAGTA
1500 1510 1520 1530 1540
320 330 340 350 360 370 380
TCTTCCCAAGGCTCTCCGAGAGTCTCCCAAGAGTCTCTAGACTCTCTCTCTCTCTCTCTCTCTCTCT
1550 1560 1570 1580 1590 1600 1610
GTCCTCCCAAGGCTCTCCCAAGAGTCTCTCCCAAGAGTCTCTAGACTCTCTCTCTCTCTCTCTCTCT
1550 1560 1570 1580 1590 1600 1610
390 400 410 420 430 440 450
CCCAGTGGATGAGGAGTCCAGAG---TGAAGAGGAGGATACAGCTACTTGGCATGCTTCCAGAGAGT
1620 1630 1640 1650 1660 1670 1680
GCTCATTCAGTGGAGTCCAGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1620 1630 1640 1650 1660 1670 1680
```

6. US-10-085-108-21 (1-1041)

US-09-066-281B-2 Sequence 20, Application US/09066281B

Initial Score = 434 Optimized Score = 704 Significance = 1.25
Residue Identity = 73% Matches = 764 Mismatches = 211
Gaps = 71 Conservative Substitutions = 0

```
1240 1250 1260 1270 1280 1290 1300
CATCCATATCCCTGTGATAGCTTTACCTGTCTCTGTAAGAGTGTCTCTCCGCTTCCAGGCTTC
1240 1250 1260 1270 1280 1290 1300
X
ATGCCCTCTCTTCCAAACCTTC
1240 1250 1260 1270 1280 1290 1300
CACCCCTCAGCTTTGAGGAAGACTTCCAGAACCCGAGTGTGACAGAGACTTGTAGATGACAGAGTCCA
1310 1320 1330 1340 1350 1360 1370 1380
CATTCGCAAGCTTTGACAAACACTCCCGGACTCAGTTGAGTTAGAGACTGGGTAGANGACAGCATCCA
1310 1320 1330 1340 1350 1360 1370 1380
100 110 120 130 140 150 160
TAGATGAGGAGGAGGAGTCCCTCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1390 1400 1410 1420 1430 1440
CAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1390 1400 1410 1420 1430 1440
170 180 190 200 210 220 230
CCTTGTCTCTATCTCTACCCCTGTGCTTACCCCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1450 1460 1470 1480 1490
TCT-----CCACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1450 1460 1470 1480 1490
240 250 260 270 280 290 300 310
ATATGCTGCTGCTGGAGTGCACCTCTTCCCAAGAGTCTCTAGACTCTCTCTCTCTCTCTCTCTCTCT
1500 1510 1520 1530 1540
AGGTGCCCTCTGCTGATACCAATCTTACCGAGA-----GCATTCCCAAGTA
1500 1510 1520 1530 1540
320 330 340 350 360 370 380
TCTTCCCAAGGCTCTCCGAGAGTCTCCCAAGAGTCTCTAGACTCTCTCTCTCTCTCTCTCTCTCTCT
1550 1560 1570 1580 1590 1600 1610
GTCCTCCCAAGGCTCTCCCAAGAGTCTCTCCCAAGAGTCTCTAGACTCTCTCTCTCTCTCTCTCTCT
1550 1560 1570 1580 1590 1600 1610
390 400 410 420 430 440 450
CCCAGTGGATGAGGAGTCCAGAG---TGAAGAGGAGGATACAGCTACTTGGCATGCTTCCAGAGAGT
1620 1630 1640 1650 1660 1670 1680
GCTCATTCAGTGGAGTCCAGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1620 1630 1640 1650 1660 1670 1680
```

7. US-10-085-108-21 (1-1041)

US-09-468-433C-2 Sequence 20, Application US/09468433C

Initial Score = 434 Optimized Score = 704 Significance = 1.25
Residue Identity = 73% Matches = 764 Mismatches = 211
Gaps = 71 Conservative Substitutions = 0

```
1030 X
CCAGTATCATCCCTAG
1030
ACAACTATGTTCTAGTTCCTTTCCATCCTGGTACAAAGATGCTTTGAAAGATGTGGAAGAGAGAG
2260 2270 2280 2290 2300 2310
CATCCATATCCCTGTGATAGCTTTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1240 1250 1260 1270 1280 1290 1300
ATGCCCTCTCTTCCAAACCTTC
1240 1250 1260 1270 1280 1290 1300
CACCCCTCAGCTTTGAGGAAGACTTCCAGAACCCGAGTGTGACAGAGACTTGTAGATGACAGAGTCCA
1310 1320 1330 1340 1350 1360 1370 1380
CATTCGCAAGCTTTGACAAACACTCCCGGACTCAGTTGAGTTAGAGACTGGGTAGATGACAGCATCCA
1310 1320 1330 1340 1350 1360 1370 1380
```


[illegible]

```

1760      1770      1780      1790      1800      1810      1820
600      610      620      630      640      650      660
TCTTCGGAAAGCCCATGAGTTTCATAGAGCTAAATTTTGGCAATTGCCCTGACTGATATGACCCCGACAACC
TACTCAAGAGAGCCCGTGAGTTCATGTGAGGCTCTTTTTGGCCTTGCCCTGATAGAAGTGGGCCCTG---ACC
1830      1840      1850      1860      1870      1880      1890
670      680      690      700      710      720      730
ACTCCTATTTCTTTGAACACACATAGACCTCACTCATGAGGGAAGCCTGANTGATGACCAAGGCATGCCCA
ACTTCTGTGTGTTTGSCAACACACAGTAGGCCCTCACCGATGAGGTAG-----TGATGATGAGGCAATGCCCG
1900      1910      1920      1930      1940      1950      1960
740      750      760      770      780      790      800      810
AGAACTGTCTCTGATTTCTTATTTCTTCAGTATGATCTTCATAAAGGGCAGCTGTCTCCCCGAGGAGTTCATCT
AGAACAGCTCTCTGATTAATTATTTCTGAGTGTGATCTTCATAAGGGCAAATGTGCGCTCTGAGGAGTCACTC
1970      1980      1990      2000      2010      2020      2030
820      830      840      850      860      870      880
GGGAAGTGTTCAGTGCAAATAGGGGTGTGTCTGGGAGGGAGCACTTTATATATGGGATCCCAGAAAGCTGC
GGGAAGTGTCTGAATGCAGTAGGAGGTATATGCTGGAGGGAGCACTTCGTCTATGGGAGCGCTTAGGAGCTCC
2040      2050      2060      2070      2080      2090      2100
890      900      910      920      930      940      950
TCACTATACATTTGGGTGCAGAGAAAGTACCTTGGAGTACCGGAGGTGCCCAACAGTGTCTCTCCACGTTATG
TCACTAAAGTTTGGGTGCAGGGAATTACTTGGAGTATCGGAGGTGCCCAAGTTCCTCCCATTTATG
2110      2120      2130      2140      2150      2160      2170
960      970      980      990      1000      1010      1020
AATTTTGTGGGTGCCAAGAGCCCATTCAGAGGCCACAAGA--GAAGTCTTAGAGTTTTTATPCCAAGCTAT
AATTCCTGTGGGTCCAAGAGCCCATTCAGAAAGCATCAGAGAGAAGTACTAGATTTTTTACCCAAGCTGA
2180      2190      2200      2210      2220      2230      2240      2250
1030      X
CCAGTATCATCCCTAG
ACAACATGTTCTTAGTTCCTTTCCATCTCTGTACAGGATGCTTTGAAGAATGTGGAAGAGAGAG
2260      2270      2280      2290      2300      2310
US-10-085-108-21 (1-1041)
US-08-845-528C-8 Sequence 8, Application US/08845528C

Initial Score = 396 Optimized Score = 561 Significance = 1.08
Residue Identity = 57% Matches = 612 Mismatches = 415
psps          = 44 Conservative Substitutions = 0

X              10       20       30       40
ATGCTCTCTTTCCAAACCTTCCAAGCTTCAGCTCAGCTTTGAGG--AGAGCTTCC
10      20 x    30      40      50      60      70
CCATCTGAGGGACGGGTAGAGTTGCGGCCGAAGAACCTTGACCCAGGCTCTGTGAGGAGGCAAGGTTTTC
50      60      70      80      90      100     110
AGAACCGAGTGTGCACAGAGGACTTGTAGATGCACAGGATTCATAGATGAGGAGGAGGAGA-----TGC
GGGACACAGGCCAACCCAGAGGAC----AGAGTTCCCTGGAGGGCCACAGAGGACCAAGAGGAGAGATCTGC
80      90      100     110     120     130     140
120      130      140      150      160      170      180      190      200      210
CTCCTCCACTTCCTTCCCTTTCACACTTTTATTCCTCCCTCTCTCTCTTCCTTCCTCATCTCCACCTT
CTGTGGGFTCTCATTGCCCAAGCTCTCGCCCAACATCTCGTGCCTGTGCTGCTGACGAGAGTCATCATGTCCTCT
150      160      170      180      190      200      210
190      200      210      220      230      240      250
GTCTCTCACCCCTTACCTCTACTCTCATCTTGGGTGTTCCAGAGATGAGGATATGCTCTGCTGGG---AT

```

[illegible]

680 690 700 710 720 730 740
CTATTCTTTTGAACACATTAGACCTCACTATGAGGAAGCTGATTGATGACCGGCGATGCCAAGAA

GCAGAGTCCTCCCGCCAGAGTCCTCTAGACTCTGCTC-----ATCCCTCTTTTCTGGAC

[illegible][illegible]

X
 1.0 2.0 3.0 4.0 5.0 6.0 7.0
 ATGCGCTCTCTTTCCAAACCTTCCACGCTCAGCTTTGAGGAGAGACTTCAGAACCCGAGTGTGACAGAGAC
 |||||
 ATGCGCTCGGGGTCAACAGAGTAAGCTCCGTACTGTGAGAAACGCGAGAGACCAATGGTTCAGCCACAGGTT
 X
 1.0 2.0 3.0 4.0 5.0 6.0 7.0

[illegible]

GCAATAGGGGTGTGCTGGGAGGAGCACATTTATATATATGGGATCCGAAAGAGCTGCTCACTATACATGG
 CTGTGGGGATATATGATGGATGCTTGCATTTCAATCTATGGGATGCTCGAAGATCATTTACTGAAGATT
 940 950 960 970 980 990 1000
 900 910 920 930 940 950 960
 GTGCAGAGAAAGTACCTTGGAGTACCGGAGGTGCCAACAGTCTCTCCACGTTATGAATTTTGTGGGGT
 GTGCAAGATTAAGTACGTTGGTTTACCGGAGGTGTGCAACAGTATCTCCATGCTATGAGTTCCTGTGGGT
 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150
 970 980 990 1000 1010 1020 1030
 CCAAGAGCCCATTCAGAGGCCCAAGA--GAAGTCTTAGAGTTTATCCAAGCTATCCAGTATCATCCCT
 CCAGAGCCCTATGCTGAACACCACCAAGATGAGAGTCTCTGCTGTTTGGCCGACAGCAGTAACACAGATCCC
 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150
 AG
 GGTTTATACCAATCTGTATGAAGACGTTGTATAGATGAGGTAGAGAG
 X 1160 1170 1180 1190 1200
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640 2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760 2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880 2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560 4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400 5410 5420 5430 5440 5450 5460 5470 5480 5490 5500 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600 5610 5620 5630 5640 5650 5660 5670 5680 5690 5700 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300 6310 6320 6330 6340 6350 6360 6370 6380 6390 6400 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600 6610 6620 6630 6640 6650 6660 6670 6680 6690 6700 6710 6720 6730 6740 6750 6760 6770 6780 6790 6800 6810 6820 6830 6840 6850 6860 6870 6880 6890 6900 6910 6920 6930 6940 6950 6960 6970 6980 6990 7000 7010 7020 7030 7040 7050 7060 7070 7080 7090 7100 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200 7210 7220 7230 7240 7250 7260 7270 7280 7290 7300 7310 7320 7330 7340 7350 7360 7370 7380 7390 7400 7410 7420 7430 7440 7450 7460 7470 7480 7490 7500 7510 7520 7530 7540 7550 7560 7570 7580 7590 7600 7610 7620 7630 7640 7650 7660 7670 7680 7690 7700 7710 7720 7730 7740 7750 7760 7770 7780 7790 7800 7810 7820 7830 7840 7850 7860 7870 7880 7890 7900 7910 7920 7930 7940 7950 7960 7970 7980 7990 8000 8010 8020 8030 8040 8050 8060 8070 8080 8090 8100 8110 8120 8130 8140 8150 8160 8170 8180 8190 8200 8210 8220

```
810 820 830 840 850 860 870
CACTCGGAGTCTTGTGATGCAATAGGGGTGTGTGCTGGAGGAGCACTTTATATATGGGATCCAGAAA
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
TACCTGGAATTTCTGGATATGATGCAAAATATATATGATGGGAGAGAGTACTACATCTATGGAGAGCCAGGAA
460 470 480 490 500 510 520
880 890 900 910 920 930 940 950
GCTGCTCACTATACATTGGGTGAGAGAAAGTACCTGGAGTACGGGAGGTGCCCAACAGTGTCTCTCCACG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
GCTCACTCACTCAGGATTTCTGTGAGGCTTACGTACCTGGAGTACCACCGGTGCCCTGCACTTATCTCTGCACA
530 540 550 560 570 580 590
960 970 980 990 1000 1010 1020
TTATGAATTTTGTGGGTCCAAGAGCCCAATTTCAGAGGCCAGCAAGA--GAAGTCTTAGAGTTTTTATCCAA
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CTATCAATTCCTTTGGGGTCCAAGAGCCTATACTGAAACCAGCAAGATGAAGTCTCTGGGAATTTTGGCCAA
600 610 620 630 640 650 660
1030 X
GCTATCCAGTATCATCCCTAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
GGTCAATGATATTGCTCCAGGTGCTTCTCATCAATATGAAGAGGCTTTGCAAGATGAGGAAGAGAGCC
670 680 690 700 710 720 730
```


CURRENT APPLICATION NUMBER: US/10/189,266
CURRENT FILING DATE: 2002-07-02
NUMBER OF SEQ ID NOS: 150
SEQ ID NO 118
LENGTH: 20
TYPE: DNA
ORGANISM: H. sapiens
FEATURE:
US-10-189-266-118

Query Match 61.0%; Score 12.8; DB 12; Length 20;
Best Local Similarity 87.5%; Pred. No. 6.7e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 TGATTGATGACCAGG 21
|||||
DB 1 TGATTGATGACAAAG 16

RESULT 3
US-10-115-223-33
Sequence 33, Application US/10115223
Publication No. US20030176334A1
GENERAL INFORMATION:
APPLICANT: Brooks, Peter
APPLICANT: Cheres, David A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR INHIBITION OF
FILE REFERENCE: MER00495
CURRENT APPLICATION NUMBER: US/10/115,223
CURRENT FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: US/09/194,468
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/018,773
PRIOR FILING DATE: 1996-05-31
PRIOR APPLICATION NUMBER: 60/015,896
PRIOR FILING DATE: 1996-05-31
PRIOR APPLICATION NUMBER: PCT/US97/09158
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 33
LENGTH: 21
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: oligonucleotide primer
US-10-115-223-33

Query Match 60.0%; Score 12.6; DB 13; Length 21;
Best Local Similarity 78.9%; Pred. No. 8.6e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GCCTGATTGATGACCAGG 21
|||||
DB 1 GCCGATCCATGACCAAGT 19

RESULT 4
US-10-251-598-93
Sequence 0, Application US/10251598
Publication No. US20030170668A1
GENERAL INFORMATION:
APPLICANT: Detera-Wadleigh, Sevilla D.
Gershon, Elliot S.
Badner, Judith A.
Goldin, Lynn R.
Berretini, Wade H.
Yoshikawa, Takeo
Sanders, Alan R.
Esterling, Lisa E.
TITLE OF INVENTION: Chromosomal Markers and Diagnostic

NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/251,598
FILING DATE: 19-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/091,952
FILING DATE: 19-Apr-1999
APPLICATION NUMBER: US 60/029,278
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: PCT/US97/19381
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-297100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: D18S1299 reverse primer
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY:
LOCATION: 1...20
SEQUENCE DESCRIPTION: SEQ ID NO: 93:
US-10-251-598-93

Query Match 59.0%; Score 12.4; DB 13; Length 20;
Best Local Similarity 92.9%; Pred. No. 1.1e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GATTGATGACCAGG 20
|||||
DB 2 GATTGAGGACCAGG 15

RESULT 5
US-09-866-108-8495
Sequence 8495, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharron G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AEOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 8495
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-8495

Query Match 58.1%; Score 12.2; DB 9; Length 17;
Best Local Similarity 82.4%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGCTGATTGATGACCA 18
||||| ||||| ||
Db 1 AGCTGGTGGATGACCA 17

RESULT 6
US-09-866-108-8495
; Sequence 8496, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:

; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 8496
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-8496

Query Match 58.1%; Score 12.2; DB 9; Length 17;
Best Local Similarity 82.4%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GCCTGATTGATGACCAG 19
||||| ||||| ||
Db 1 GCCTGGTGGATGAGCAG 17

RESULT 7
US-10-168-989-15/c
; Sequence 15, Application US/10168989
; Publication No. US20030190631A1
; GENERAL INFORMATION:
; APPLICANT: Chartier-Harlin et al.
; TITLE OF INVENTION: Implication of a known gene named CP2/LSF-LBP-1 in
; FILE REFERENCE: P07666US00/BAS
; CURRENT APPLICATION NUMBER: US/10/168,989
; CURRENT FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-168-989-15

Query Match 57.1%; Score 12; DB 13; Length 20;
Best Local Similarity 75.0%; Pred. No. 1.8e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AGCTGATTGATGACCAGG 21
||||| ||||| ||
Db 20 AGCCAGATTACAGCAGG 1

RESULT 8
US-10-090-011-60/c
; Sequence 60, Application US/10090011
; Publication No. US20030082810A1
; GENERAL INFORMATION:
; APPLICANT: Serup, Palle
; APPLICANT: Heimberg, Harry
; APPLICANT: Gradwohl, Gerard
; TITLE OF INVENTION: Methods For Generating Insulin-Secreting
; TITLE OF INVENTION: Cells Suitable for Transplantation

Local Similarity: 80.7%, Freq. NO: 2.2E+04;
 Matches 13; Conservative 0; Mismatches 2;
 Indels 0; Gaps 0;

QV 2 AGCCTGATTGATGACCAG 19

Db 2 AGCCGAGTGTCTGACCAG 19
|||||

RESULT 13

US-10-159-856-41/c
; Sequence 41, Application US/10159856
; Publication No. US20030228689A1
; GENERAL INFORMATION:
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF G PROTEIN-COUPLED RECEPTOR KINASE 6 EXPRESSION
; FILE REFERENCE: RTS-0365
; CURRENT APPLICATION NUMBER: US/10/159,856
; CURRENT FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 134
; SEQ ID NO 41
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-159-856-41

Query Match 55.2%; Score 11.6; DB 12; Length 20;
Best Local Similarity 77.8%; Pred. No. 2.9e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 CCTGATTGATGACACGAGG 21
|||
Db 19 CTTGCTGGATGACACGAG 2

RESULT 14

US-10-159-856-106
; Sequence 106, Application US/10159856
; Publication No. US20030228689A1
; GENERAL INFORMATION:
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF G PROTEIN-COUPLED RECEPTOR KINASE 6 EXPRESSION
; FILE REFERENCE: RTS-0365
; CURRENT APPLICATION NUMBER: US/10/159,856
; CURRENT FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 134
; SEQ ID NO 106
; LENGTH: 20
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
US-10-159-856-106

Query Match 55.2%; Score 11.6; DB 12; Length 20;
Best Local Similarity 77.8%; Pred. No. 2.9e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 CCTGATTGATGACACGAGG 21
|||
Db 2 CTTGCTGGATGACACGAG 19

RESULT 15

US-09-382-860-144
; Sequence 144, Application US/09382860
; Publication No. US20030110526A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Jr., Robert H.
; APPLICANT: Liu, Jing
; APPLICANT: Aoki, Masashi
; APPLICANT: Hoffman, Eric
; APPLICANT: Chou, Fan-Li
; TITLE OF INVENTION: DYSFERLIN MUTATIONS
; FILE REFERENCE: 00786/401002

; CURRENT APPLICATION NUMBER: US/09/382,860
; CURRENT FILING DATE: 1999-08-25
; EARLIER APPLICATION NUMBER: US 60/097,930
; EARLIER FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 144
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-382-860-144

Query Match 55.2%; Score 11.6; DB 11; Length 21;
Best Local Similarity 77.8%; Pred. No. 2.9e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GCCTGATTGATGACACGAG 20
|||
Db 4 GCCTGAGGATCAGCAGG 21

Search completed: February 20, 2004, 04:09:31
Job time : 198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 00:42:09 ; Search time 45.5 seconds
(without alignments)
203.715 Million cell updates/sec

Title: US-10-085-108-21_COPY_711_731

Perfect score: 21
Sequence: 1 AAGCTGATGATGACAGGG 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 352324

Minimum DB seq length: 0
Maximum DB seq length: 21

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12.8	61.0	19	1	US-08-200-232-16
2	12.8	61.0	19	5	PCT-US95-02219-16
3	12.8	61.0	19	5	PCT-US95-02219A-17
4	12.6	60.0	21	4	US-09-194-468A-33
5	12.4	59.0	15	1	US-08-363-240A-560
6	12.4	59.0	19	3	US-09-050-159-38
7	12.4	59.0	20	4	US-09-091-952A-93
8	12.2	58.1	20	4	US-09-658-679A-61
9	11.8	56.2	21	1	US-08-171-718-7
10	11.8	56.2	21	3	US-08-478-087-7
11	11.8	56.2	21	4	US-09-081-149-17
12	11.6	55.2	19	1	US-08-363-233B-7
13	11.6	55.2	20	2	US-08-809-297-30
14	11.6	55.2	20	4	US-09-554-511-20
15	11.6	55.2	21	3	US-08-874-825-35
16	11.6	55.2	21	3	US-08-954-536-7
17	11.6	55.2	21	3	US-08-954-536-7
18	11.6	55.2	21	3	US-09-387-699-22
19	11.6	55.2	21	4	US-09-231-303-35
20	11.6	55.2	21	4	US-09-641-259B-22
21	11.4	54.3	15	1	US-08-291-932A-31
22	11.4	54.3	15	1	US-08-291-932A-220
23	11.4	54.3	15	1	US-08-363-240A-561
24	11.4	54.3	15	1	US-08-363-240A-562
25	11.4	54.3	20	4	US-09-954-560-37
26	11.2	53.3	18	2	US-08-284-465-11
27	11.2	53.3	18	4	US-09-580-189-12

28	11.2	53.3	19	4	US-09-422-978-4273	Sequence 4273, Ap
29	11.2	53.3	19	4	US-09-230-652-111	Sequence 111, Ap
30	11.2	53.3	20	1	US-08-044-618-2	Sequence 2, Appl
31	11.2	53.3	20	2	US-08-334-545-8	Sequence 8, Appl
32	11.2	53.3	20	3	US-08-827-036A-13	Sequence 13, Appl
33	11.2	53.3	20	3	US-09-147-933-48	Sequence 48, Appl
34	11.2	53.3	20	4	US-09-467-642-89	Sequence 89, Appl
35	11.2	53.3	20	4	US-09-383-316-110	Sequence 110, Appl
36	11.2	53.3	20	4	US-09-792-594-14	Sequence 14, Appl
37	11.2	53.3	20	4	US-09-198-452A-6317	Sequence 6317, Ap
38	11.2	53.3	21	2	US-08-555-723B-8	Sequence 8, Appl
39	11.2	53.3	21	3	US-09-123-465-8	Sequence 8, Appl
40	11	52.4	19	1	US-07-977-284A-21	Sequence 21, Appl
41	11	52.4	19	2	US-08-256-426B-21	Sequence 21, Appl
42	11	52.4	20	4	US-09-596-938-9	Sequence 9, Appl
43	11	52.4	21	2	US-08-690-734A-44	Sequence 44, Appl
44	11	52.4	21	3	US-08-742-185-44	Sequence 44, Appl
45	11	52.4	21	4	US-09-305-927-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-08-200-232-16
; Sequence 16, Application US/08200232
; Patent No. 5721349
; GENERAL INFORMATION:
; APPLICANT: Cover, Timothy L.
; APPLICANT: Blaser, Martin J.
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI
; TITLE OF INVENTION: AND RELATED METHODS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/200,232
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 2200.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-200-232-16

Query Match 61.0%; Score 12.8; DB 1; Length 19;
Best Local Similarity 87.5%; Pred. No. 9.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAGCTGATGATGAC 16

Db 1 AAGCTGATGATGAC 16

RESULT 2
PCT-US95-02219-16
; Sequence 16, Application PC/TUS9502219
; GENERAL INFORMATION:
; APPLICANT: Cover, Timothy L.
; APPLICANT: Blaser, Martin J.
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI
; TITLE OF INVENTION: AND RELATED METHODS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02219
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 2200.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-9880
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US95-02219-16

Query Match 61.0%; Score 12.8; DB 5; Length 19;
Best Local Similarity 87.5%; Pred. No. 9.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGAC 16
Db 1 AAGCTGATTGATGAC 16

RESULT 3
PCT-US95-02219A-17
; Sequence 17, Application PC/TUS9502219A
; GENERAL INFORMATION:
; APPLICANT: Cover, Timothy L.
; APPLICANT: Tummuru, Murali KR
; APPLICANT: Cao, Ping
; APPLICANT: Thompson, Stuart A.
; APPLICANT: Blaser, Martin J.
; TITLE OF INVENTION: VACUOLATING TOXIN-DEFICIENT H. PYLORI
; TITLE OF INVENTION: AND THE RELATED METHODS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02219A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 2200.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US95-02219A-17

Query Match 61.0%; Score 12.8; DB 5; Length 19;
Best Local Similarity 87.5%; Pred. No. 9.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGAC 16
Db 1 AAGCTGATTGATGAC 16

RESULT 4
US-09-194-468A-33
; Sequence 33, Application US/09194468A
; Patent No. 650924
; GENERAL INFORMATION:
; APPLICANT: Brooks, Peter
; APPLICANT: Cheres, David A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR INHIBITION OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: MER00498
; CURRENT APPLICATION NUMBER: US/09/194,468A
; CURRENT FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/018,773
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: 60/015,896
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: PCT/US97/09158
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide primer
US-09-194-468A-33

Query Match 60.0%; Score 12.6; DB 4; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GCCTGATTGATGACAGG 21
Db 1 GCCGGATCCATGACCATG 19

RESULT 5
US-08-363-240A-560/c
; Sequence 560, Application US/08363240A
; Patent No. 5705388

;; GENERAL INFORMATION:
;; APPLICANT: Couture, Larry
;; APPLICANT: McSwiggen, James
;; APPLICANT: Bisgaier, Charles
;; APPLICANT: Pape, Michael
;; TITLE OF INVENTION: METHOD AND REAGENT FOR
;; TITLE OF INVENTION: PREVENTION, INHIBITION OF
;; TITLE OF INVENTION: PROGRESSION AND REGRESSION
;; TITLE OF INVENTION: OF VASCULAR DISEASES
;; NUMBER OF SEQUENCES: 1243
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Lyon & Lyon
;; STREET: 633 West Fifth Street
;; STREET: Suite 4700
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90071
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: IBM P.C. DOS 5.0
;; SOFTWARE: Word Perfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/363,240A
;; FILING DATE: December 23, 1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warburg, Richard
;; REGISTRATION NUMBER: 32,327
;; REFERENCE/DOCKET NUMBER: 210/036
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;; INFORMATION FOR SEQ ID NO: 560:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-363-240A-560

Query Match 59.0%; Score 12.4; DB 1; Length 15;
Best Local Similarity 92.9%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CTGATTGATGACCA 18
Db 14 CTGATTGATGACCA 1

RESULT 6
US-09-050-159-38
; Sequence 38, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Leif T
; APPLICANT: Andersson, Maria K
; APPLICANT: Linstrom, Per H
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; TITLE OF INVENTION: COMPOSITIONS FOR USE THEREOF
; FILE REFERENCE: 1248/1D042
; CURRENT APPLICATION NUMBER: US/09/050,159A
; CURRENT FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042,930
; EARLIER FILING DATE: 1987-04-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38

;; LENGTH: 19
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: PCR PRIMER
US-09-050-159-38

Query Match 59.0%; Score 12.4; DB 3; Length 19;
Best Local Similarity 92.9%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGCCTGATTGATGA 15
Db 4 AGCCTGATTGATGA 17

RESULT 7
US-09-091-952A-93
; Sequence 93, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; Gershon, Elliot S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berrettini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1...20
; OTHER INFORMATION: D18S1299 reverse primer
; SEQUENCE DESCRIPTION: SEQ ID NO: 93:

US-09-091-952A-93

Query Match 59.0%; Score 12.4; DB 4; Length 20;
Best Local Similarity 92.9%; Pred. No. 1.5e-03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GATTGATGACGAGG 20
|||||
Db 2 GATTGAGGACGAGG 15

RESULT 8

US-09-658-679A-61/C
; Sequence 61, Application US/09658679A
; Patent No. 644464
; GENERAL INFORMATION:
; APPLICANT: Ian Popoff
; APPLICANT: Jacqueline Wvatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF E2F TRANSCRIPTION FACTOR 2 EXPRESSION
; FILE REFERENCE: RTS-0186
; CURRENT APPLICATION NUMBER: US/09/658,679A
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 61
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-658-679A-61

Query Match 58.1%; Score 12.2; DB 4; Length 20;
Best Local Similarity 82.4%; Pred. No. 1.9e-03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCTGATTGATGACGAGG 20
|||||
Db 18 CCTGACTGAGGACGAGG 2

RESULT 9

US-08-171-718-7
; Sequence 7, Application US/08171718
; Patent No. 5707863
; GENERAL INFORMATION:
; APPLICANT: Trofatter, James A.
; APPLICANT: MacCollin, Mia M.
; APPLICANT: Gusella, James F.
; TITLE OF INVENTION: Tumor Suppressor Gene Merlin and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/171,718
; FILING DATE: 22-DEC-1993
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/108,808
; FILING DATE: 19-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/022,034
; FILING DATE: 25-FEB-1993
; INFORMATION FOR SEQ ID NO: 7:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/026,063
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 0609.3850003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-171-718-7

Query Match 56.2%; Score 11.8; DB 1; Length 21;
Best Local Similarity 86.7%; Pred. No. 3.1e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCCTGATTGATGAC 16
|||||
Db 5 AACCTGATTGATGAC 19

RESULT 10

US-08-478-087-7
; Sequence 7, Application US/08478087
; Patent No. 6077685
; GENERAL INFORMATION:
; APPLICANT: Trofatter, James A.
; APPLICANT: MacCollin, Mia M.
; APPLICANT: Gusella, James F.
; TITLE OF INVENTION: Tumor Suppressor Gene Merlin and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,087
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,718
; FILING DATE: 22-DEC-1993
; APPLICATION NUMBER: US 08/108,808
; FILING DATE: 19-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/022,034
; FILING DATE: 25-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/026,063
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 0609.3850003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 7:

CLASSIFICATION: 435

```
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "SYNTHETIC DNA"
US-08-809-297-30

Query Match          55.2%; Score 11.6; DB 2; Length 20;
Best Local Similarity 77.8%; Pred. No. 3.9e+03;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  2 AGCTGATTGATGACCAG 19
Db  18 AGCATGATCATGACCTG 1

RESULT 14
US-09-554-511-20
; Sequence 20, Application US/09554511
; Patent No. 6312928
; GENERAL INFORMATION:
; APPLICANT: VAN GEMEN, Bob
; APPLICANT: VAN STRIUP, Dianne A.M.
; APPLICANT: SCHUKKINK, Adriana F.
; TITLE OF INVENTION: TRANSCRIPTION BASED AMPLIFICATION OF DOUBLE STRANDED
; FILE REFERENCE: VAN GEMEN ET AL.
; CURRENT APPLICATION NUMBER: US/09/554,511
; CURRENT FILING DATE: 2000-05-16
; PRIOR APPLICATION NUMBER: PCT/EP98/07329
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: EP 97203577.8
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 20
; LENGTH: 20
; TYPE: DNA
; ORGANISM: HUMAN
US-09-554-511-20

Query Match          55.2%; Score 11.6; DB 4; Length 20;
Best Local Similarity 77.8%; Pred. No. 3.9e+03;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  3 GCCTGATTGATGACCAG 20
Db  1 GCTTGATGATGACCATG 18

RESULT 15
US-08-874-825-35
; Sequence 35, Application US/08874825
; Patent No. 6057101
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Rothberg, Jonathan
; APPLICANT: Yang, Meijia
; APPLICANT: Knight, James
; APPLICANT: Kalbfleisch, Theodore
; TITLE OF INVENTION: IDENTIFICATION AND COMPARISON OF
; TITLE OF INVENTION: PROTEIN-PROTEIN INTERACTIONS THAT OCCUR IN POPULATIONS
; TITLE OF INVENTION: AND IDENTIFICATION OF INHIBITORS OF THESE INTERACTORS
; NUMBER OF SEQUENCES: 122
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
```

```
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,825
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/663,824
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
US-08-874-825-35

Query Match          55.2%; Score 11.6; DB 3; Length 21;
Best Local Similarity 66.7%; Pred. No. 4e+03;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy  4 CCTGATTGATGACCGGG 21
Db  2 CCUGAUGGUGAGACCGGG 19
```

Search completed: February 20, 2004, 02:40:59
Job time : 46.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 00:32:29 ; Search time 1490 Seconds
(without alignments)
342.546 Million cell updates/sec

Title: US-10-085-108-21_COPY_711_731

Perfect score: 21

Sequence: 1 AAGCTGATTGATGACGAGG 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 8380

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estim.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	10.2	48.6	20	28	AZ600911
C 2	10.2	48.6	21	28	AZ306912
C 3	9.8	46.7	19	28	AZ763729
C 4	9.8	46.7	21	28	AZ504285

C 5	9.4	44.8	19	14	C01992
C 6	9.4	44.8	21	28	AZ771145
C 7	9.2	43.8	21	28	AZ509602
C 8	9.2	43.8	21	28	AZ661528
C 9	9.2	43.8	21	28	AZ662959
C 10	9.2	43.8	21	28	AZ774704
C 11	8.8	41.9	16	12	BM395054
C 12	8.8	41.9	19	14	C01993
C 13	8.8	41.9	21	28	AZ987077
C 14	8.6	41.0	18	13	BQ594466
C 15	8.6	41.0	20	28	AZ307610
C 16	8.6	41.0	20	28	AZ462631
C 17	8.6	41.0	20	28	AZ665083
C 18	8.6	41.0	21	28	AZ405406
C 19	8.4	40.0	15	14	CA796369
C 20	8.4	40.0	19	9	AW249918
C 21	8.4	40.0	19	28	AZ808212
C 22	8.4	40.0	20	28	AZ308421
C 23	8.4	40.0	20	28	AZ473322
C 24	8.2	39.0	19	28	AZ491924
C 25	8.2	39.0	19	28	AZ655467
C 26	8.2	39.0	19	28	AZ655870
C 27	8.2	39.0	19	28	AZ815827
C 28	8.2	39.0	19	28	AZ875430
C 29	8.2	39.0	19	28	AZ937956
C 30	8.2	39.0	19	28	AZ989459
C 31	8.2	39.0	19	28	AZ990851
C 32	8.2	39.0	20	29	TA345E06Q
C 33	8.2	39.0	21	9	AW332721
C 34	8.2	39.0	21	28	AZ796024
C 35	8.2	39.0	21	28	AZ806895
C 36	8	38.1	19	28	AZ303949
C 37	8	38.1	19	28	AZ810098
C 38	8	38.1	19	28	AZ836630
C 39	8	38.1	19	28	AZ849303
C 40	8	38.1	20	9	AU254575
C 41	8	38.1	20	13	BQ583464
C 42	8	38.1	21	28	AZ775019
C 43	7.8	37.1	19	12	BM396288
C 44	7.8	37.1	19	28	AZ486389
C 45	7.8	37.1	19	28	AZ621254

ALIGNMENTS

RESULT 1
AZ600911/c

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ600911 20 bp DNA linear GSS 13-DEC-2000
1M0418M21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0418M21 R, genomic survey sequence.
AZ600911 GI:11723101
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islan, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0418 row: M column: 21
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 20.
 Location/Qualifiers

FEATURES

source

```

1. .20
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUGC1M0418M21"
  /sex="Male"
  /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
  /clone_lib="Mouse 10kb plasmid UUGC1M library"
  /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
  2 a 7 c 6 g 5 t

```

BASE COUNT

ORIGIN

```

Query Match      48.6%; Score 10.2; DB 28; Length 20;
Best Local Similarity 80.0%; Pred. No. 5.3e+05;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 6 TGATTGATGACACAGG 20
    ||||| |||||
DB 18 TGATGCCGACACAGG 4

```

RESULT 2

AZ306912

LOCUS

DEFINITION

1M0008L07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0008L07 F, genomic survey sequence.

ACCESSION

AZ306912

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

AZ306912      21 bp  DNA  linear  GSS 29-SEP-2000
1M0008L07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0008L07 F, genomic survey sequence.
AZ306912.1  GI:10345389
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 21)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606

```

Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0008 row: L column: 07
 Seq primer: CGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 21.
 Location/Qualifiers

FEATURES

source

```

1. .21
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUGC1M0008L07"
  /sex="Male"
  /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
  /clone_lib="Mouse 10kb plasmid UUGC1M library"
  /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
  7 a 2 c 6 g 6 t

```

BASE COUNT

ORIGIN

```

Query Match      48.6%; Score 10.2; DB 28; Length 21;
Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 6 TGATTGATGACACAGG 20
    ||||| |||||
DB 1 TGATTATGCCGAGG 15

```

RESULT 3

AZ763729

LOCUS

DEFINITION

1M0559N16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0559N16 F, genomic survey sequence.

ACCESSION

AZ763729

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

AZ763729      19 bp  DNA  linear  GSS 16-FEB-2001
1M0559N16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0559N16 F, genomic survey sequence.
AZ763729      1  GI:12875056
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 19)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

```

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0559 row: N column: 16
 Seq primer: CGTGTGAACAGCAGCCAGT
 Class: plasmid ends
 High quality sequence stop: 19.
 Location/Qualifiers

FEATURES

source

1. .19
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0559N16"
 /sex="Male"
 /lab_hosts="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN

Query Match 46.7%; Score 9.8; DB 28; Length 19;
 Best Local Similarity 84.6%; Pred. No. 8e+05;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCCTGATGAT 13
 |||||
 Db 5 AAGCATGACTGAT 17

RESULT 4

AZ504285/c

LOCUS

DEFINITION 1M0344J20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0344J20 F, genomic survey sequence.

ACCESSION

AZ504285

VERSION

GSS.

KEYWORDS

SOURCE

ORGANISM

Mus musculus

(house mouse)

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 21)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0344 row: J column: 20

Seq primer: CGTGTGAACAGCAGCCAGT

Class: plasmid ends

High quality sequence stop: 21.

Location/Qualifiers

1. .21

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0344J20"

/sex="Male"

/lab_hosts="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN

Query Match 46.7%; Score 9.8; DB 28; Length 21;
 Best Local Similarity 84.6%; Pred. No. 8.4e+05;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 GATTGATGACCAG 19
 |||||
 Db 21 GATTGATCATCAG 9

RESULT 5

C01992/c

LOCUS

DEFINITION

HUMS0004015 Human adult (K.Okubo) Homo sapiens cdna, mRNA

sequence.

ACCESSION

C01992

VERSION

C01992.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 19)

Okubo,K.

BodyMap; human gene expression database

Unpublished

Contact: Okubo,K.

Institute for Molecular and Cellular Biol

Osaka University

1-3 Yamada-oka, Suita, Osaka Pref. 565, Japan

Tel: 06-877-5111(ex.3315)

Email: kousaku@imcb.osaka-u.ac.jp

We are not submitting the same cdna sequence redundantly to DDBJ

since 1993. For the abundance information of clones with this sequence in this library and as well as in other 3'-directed libraries, see <http://www.imcb.osaka-u.ac.jp/bodymap/>. The sequences of the clones represented by this GS sequences is also found there.

FEATURES

source Location/Qualifiers

1. 19
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="Human adult (K.Okubo)"
/notes="One or more human adult tissue"
6 a 5 c 4 g 4 t

BASE COUNT

ORIGIN

Query Match 44.8%; Score 9.4; DB 14; Length 19;
Best Local Similarity 90.9%; Pred. No. 1.2e+06;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCCTGATTGAT 13

Db 12 GCCTGATTGAT 2

RESULT 6

AZ771145/c

LOCUS

DEFINITION 21 bp DNA linear GSS 16-FEB-2001
clone UUGC1M0573A09 F, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0573 row: A column: 09
Seq primer: CGTTGTAAACAGCGCCAGT
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers
1. 21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0573A09"
/sex="Male"
/lab_host="F. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a

FEATURES

source

1. 21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0573A09"
/sex="Male"
/lab_host="F. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 6 c 3 g 10 t
ORIGIN

Query Match 44.8%; Score 9.4; DB 28; Length 21;
Best Local Similarity 90.9%; Pred. No. 1.3e+06;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGCTGATTGA 12

Db 11 AGCTGATTGA 1

RESULT 7

AZ509602/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0352 row: E column: 08
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers
1. 21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0352E08"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA

FEATURES

source

1. 21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0352E08"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 6 a 4 c 5 g 6 t
ORIGIN

Query Match 43.8%; Score 9.2; DB 28; Length 21;
Best Local Similarity 78.6%; Pred. No. 1.6e+06;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 ATTGATGACCAAGG 21
||||| |||||
Db 18 ATTCAACACCAAGG 5

RESULT 8
AZ661528
LOCUS
DEFINITION
1M0540106F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0540106 F, genomic survey sequence.

ACCESSION
AZ661528
VERSION
KEYWORDS
SOURCE
GSS.
Mus musculus (house mouse)

ORGANISM

REFERENCE
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.,
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
COMMENT
Unpublished
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0540 row: 1 column: 06
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers

FEATURES
source

1. .21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0540106"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 5 a 4 c 6 g 6 t
ORIGIN

Query Match 43.8%; Score 9.2; DB 28; Length 21;
Best Local Similarity 78.6%; Pred. No. 1.6e+06;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CCTGATGATGACC 17
||||| |||||
Db 4 CTTAATGATGACC 17

RESULT 9

AZ662959

LOCUS

DEFINITION

1M0542P18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0542P18 F, genomic survey sequence.

ACCESSION

AZ662959

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

1 (bases 1 to 21)

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0542 row: P column: 18

Seq primer: CGTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 21.

Location/Qualifiers

1. .21

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0542P18"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (G[14732114]gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 7 a 5 c 2 g 7 t
ORIGIN

Query Match 43.8%; Score 9.2; DB 28; Length 21;
Best Local Similarity 78.6%; Pred. No. 1.6e+06;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 CTGATTCATGACCA 18
|||||
Db 3 CTTATCATGACTA 16

RESULT 10
AZ774704
LOCUS
DEFINITION
2M0004G18F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGC2M0004G18 F, genomic survey sequence.

ACCESSION
AZ774704

VERSION
GSS.

KEYWORDS
GSS.

SOURCE
Mus musculus (house mouse)

ORGANISM

REFERENCE
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0004 row: G column: 18

Seq primer: CGTTGTAACACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 21.

Location/Qualifiers

1..21

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0004G18"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (G[14732114]gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 8 a 6 c 2 g 5 t
ORIGIN

Query Match 43.8%; Score 9.2; DB 28; Length 21;
Best Local Similarity 78.6%; Pred. No. 1.6e+06;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGCCTGATTCATCA 15
|||||
Db 1 AACCTCCTTGATCA 14

RESULT 11

BM395054/c

LOCUS

DEFINITION

50072-2-7-C10.f.2 Chilcoat/turkewitz cDNA (large fraction)

Tetrahymena thermophila cDNA, mRNA sequence.

ACCESSION

BM395054

VERSION

EST.

KEYWORDS

EST.

ORGANISM

Tetrahymena thermophila

Tetrahymena thermophila

Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;

Rhynchosomatida; Tetrahymenina; Tetrahymena.

REFERENCE

AUTHORS

Turkewitz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel,
J. and Klobutcher,L.

EST from Tetrahymena thermophila, strain CU428.1, growing cells

Unpublished

JOURNAL

COMMENT

Contact: Turkewitz AP

Molecular Genetics and Cell Biology

University of Chicago

920 E. 58th Street, Chicago, IL 60637, USA

Tel: 773 702 4374

Fax: 773 702 3172

Email: apturkew@midway.uchicago.edu

Seq primer: T3.

Location/Qualifiers

1..16

/organism="Tetrahymena thermophila"

/mol_type="mRNA"

/strain="CU428.1"

/db_xref="taxon:5911"

/clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"

/note="Vector: BlueScript2 SK+; Details on library

preparation can be found in Chilcoat and Turkewitz (2001)

Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT 5 a 5 c 4 g 2 t
ORIGIN

Query Match 41.9%; Score 8.8; DB 12; Length 16;
Best Local Similarity 83.3%; Pred. No. 2.2e+06;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GCCTGATTCATG 14


```

Db          14 GCCTGATGATG 3

RESULT 12
LOCUS      C01993/c
DEFINITION HUMGS0004016 Human adult (K.Okubo) Homo sapiens cDNA, mRNA EST 31-DEC-2002
sequence.
ACCESSION C01993
VERSION   C01993.1 GI:1434223
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 19)
AUTHORS  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
JOURNAL   Okubo,K.
COMMENT   BodyMap; human gene expression database
          Contact: Okubo,K.
          Institute for Molecular and Cellular Biol
          Osaka University
          1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan
          Tel: 06-877-5111(ex.3315)
          Email: kousaku@imcb.osaka-u.ac.jp
          We are not submitting the same cDNA sequence redundantly to DDBJ
          since 1993. For the abundance information of clones with this
          sequence in this library and as well as in other 3'-directed
          libraries, see : http://www.imcb.osaka-u.ac.jp/bodymap/. The
          sequences of the clones represented by this GS sequences is also
          found there.

FEATURES             Location/Qualifiers
     source           1..19
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="adult"
                     /clone_lib="Human adult (K.Okubo)"
                     /note="One or more human adult tissue"
BASE COUNT        6 a 5 c 4 g 3 t 1 others
ORIGIN
1
17 ACCCTGTTGNTG 5

Query Match      41.9%; Score 8.6; DB 14; Length 19;
Best Local Similarity 76.9%; Pred. No. 2.4e+06;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ACCTGATGATG 14
    |||||
Db 17 ACCCTGTTGNTG 5

RESULT 13
LOCUS      AZ987077/c
DEFINITION 2M0269H04R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0269H04 R, genomic survey sequence.
ACCESSION  AZ987077
VERSION     AZ987077.1 GI:13858504
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE 1 (bases 1 to 21)
AUTHORS  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
          Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
          M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
          and Wright,D.,Weiss,R.
          Mouse whole genome scaffolding with paired end reads from 10kb
          plasmid inserts
          Unpublished
          Contact: Robert B. Weiss

RESULT 14
LOCUS      BQ594466
DEFINITION E012442-024-024-J18-SP6 MP1Z-ADIS-024-developing root Beta vulgaris
cDNA clone 024-024-J18 5-PRIME, mRNA sequence.
ACCESSION  BQ594466
VERSION     BQ594466.1 GI:26124049
KEYWORDS    EST.
SOURCE      Beta vulgaris
ORGANISM    Beta vulgaris
REFERENCE 1 (bases 1 to 18)
AUTHORS  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Caryophyllales; Amaranthaceae; Beta.
          Herwig,R., Schulz,B., Weisshaar,B., Hennig,S., Steinfath,M.,
          Drungowski,M., Stahl,D., Wruck,W., Menze,A., O'Brien,J., Lehrach,H.
          and Radeloff,U.
          Construction of a 'unigene' cDNA clone set by oligonucleotide
          fingerprinting allows access to 25 000 potential sugar beet genes
          Plant J. 32 (5), 845-857 (2002)

```

```

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0269 row: H column: 04
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers
1..21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0269H04"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT      5 a 3 c 5 g 8 t
ORIGIN
1
8 ATTGATGACCAG 19
|||
21 ATTGATGCCAG 10

Query Match      41.9%; Score 8.8; DB 28; Length 21;
Best Local Similarity 83.3%; Pred. No. 2.5e+06;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 ATTGATGACCAG 19
    |||
Db 21 ATTGATGCCAG 10

RESULT 14
LOCUS      BQ594466
DEFINITION E012442-024-024-J18-SP6 MP1Z-ADIS-024-developing root Beta vulgaris
cDNA clone 024-024-J18 5-PRIME, mRNA sequence.
ACCESSION  BQ594466
VERSION     BQ594466.1 GI:26124049
KEYWORDS    EST.
SOURCE      Beta vulgaris
ORGANISM    Beta vulgaris
REFERENCE 1 (bases 1 to 18)
AUTHORS  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Caryophyllales; Amaranthaceae; Beta.
          Herwig,R., Schulz,B., Weisshaar,B., Hennig,S., Steinfath,M.,
          Drungowski,M., Stahl,D., Wruck,W., Menze,A., O'Brien,J., Lehrach,H.
          and Radeloff,U.
          Construction of a 'unigene' cDNA clone set by oligonucleotide
          fingerprinting allows access to 25 000 potential sugar beet genes
          Plant J. 32 (5), 845-857 (2002)

```

COMMENT

Contact: Weishaar B
ADIS DNA core facility at MPfZ
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weishaar@mpfz-koeln.mpg.de
Insert length: 18 Std Error: 0.00
Plate: 24 row: I column: 18
Seq primer: SP6; CATACGATTGATGACACTATAG.

FEATURES

source
1. 18
/organism="Beta vulgaris"
/mol_type="mRNA"
/cultivar="KWS22320 (double haploid, monogerm breeding line)"
/db_xref="taxon:161934"
/clone="024-024-118"
/tissue_type="developing root"
/lab_host="EMDH10B"
/clone_lib="MPIZ-ADIS-024-developing root"
/notes="Vector: pCMVSPORT6; Site 1: Sali; Site 2: NotI;
cDNA library from sugar beet library provided by KWS
Kleinwanzlebener Saatucht AG Einbeck, Germany, contact:
b.schulz@kws.de; cloning sites Sali-NotI, primer sites and
orientation:
SP6-Sali-CCACGGCTCCG-5prime-cDNA-polyA-CC-NotI-T7; Note:
Sequencing granted in the context of the GABI-Best project
local PI: Dr. Katharina Schneider, coordinator: Prof.
Christian Jung; Sequence submission managed by
RZPD/GABI-Primary database: http://gabi.rzpd.de"

BASE COUNT

3 a 3 c 9 g 3 t

ORIGIN

Query Match 41.0%; Score 8.6; DB 13; Length 18;
Best Local Similarity 73.3%; Pred. No. 2.9e+06;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAGCTGTGATTGATGA 15

Db 4 ACGCGTGGGTGATGA 18

RESULT 15

AZ307610

LOCUS

1M0009G15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0009G15 R, genomic survey sequence.

ACCESSION

AZ307610

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112 USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0009 row: G column: 15

Seq primer: CACACAGGAACAGCTATGACC

Class: Plasmid ends

High quality sequence stop: 20.

FEATURES

source
1. 20
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0009G15"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: pWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 4 a 6 c 4 g 6 t

ORIGIN

Query Match 41.0%; Score 8.6; DB 28; Length 20;
Best Local Similarity 73.3%; Pred. No. 3e+06;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 GATTGATGACACGGG 21

Db 1 GATTCTGCACCATG 15

Search completed: February 20, 2004, 02:39:22

Job time : 1493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:14:43 ; Search time 171 Seconds
(without alignments)
331.510 Million cell updates/sec

Title: US-10-085-108-21_COPY_711_731

Perfect score: 21

Sequence: 1 AAGCCTGATTGATGACCAAGG 21

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 1588498

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*
1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	100.0	21	25	Human MAGE-C3 expr
C 2	12.8	61.0	19	16	H.pylori tox (+/-)
C 3	12.8	61.0	19	21	cdk3 ribozyme bind
C 4	12.8	61.0	19	21	cdk3 ribozyme bind
C 5	12.8	61.0	19	22	Cell-cycle depende
C 6	12.8	61.0	19	22	Cell-cycle depende
C 7	12.6	60.0	20	16	Primer B (Group 8,
C 8	12.6	60.0	21	19	Chicken matrix met

C 9	12.6	60.0	21	19	AAV12875	5' PCR primer used
C 10	12.6	60.0	21	22	AAF96809	Human gene single
C 11	12.4	59.0	15	17	AACT50147	Rabbit CERP HH rib
C 12	12.4	59.0	18	21	AACT50147	RRV PCR primer vmi
C 13	12.4	59.0	19	19	AAV08610	Primer ACE/188RT f
C 14	12.4	59.0	19	21	AACT61238	Primer ACE, AGT and
C 15	12.4	59.0	19	21	AAA38238	Human angiotensin-
C 16	12.4	59.0	20	19	AAV57869	Human chromosome 1
C 17	12.4	59.0	21	25	AAV57869	HGF RT-PCR primer
C 18	12.2	58.1	17	24	ABN08503	Human GDMLP-1 17-m
C 19	12.2	58.1	17	24	ABN08503	Human GDMLP-1 17-m
C 20	12.2	58.1	17	25	ABT38386	Tumour suppression
C 21	12.2	58.1	20	19	AAV57185	Human Notch-3 muta
C 22	12.2	58.1	20	24	AAAD34914	Human E2F transcri
C 23	12.2	58.1	21	15	AAQ73414	Calcitonin recepto
C 24	12.2	58.1	21	25	AAAD33380	Mouse bmf DNA spec
C 25	12.2	58.1	18	24	ABX15383	Human glyceraldehy
C 26	12.2	58.1	20	22	AAA91587	PCR primer for Hum
C 27	12.2	58.1	21	24	ABK82216	Human ATP-binding
C 28	12.2	58.1	21	25	AAAD47337	Human RT-PCR rever
C 29	11.8	56.2	18	21	AAA86678	Cdc 2 kinase hamme
C 30	11.8	56.2	18	22	AAH61844	Cdc 2 kinase hamme
C 31	11.8	56.2	20	19	AAV20747	Human squalene epo
C 32	11.8	56.2	20	22	ABA82346	Znaxi gene region
C 33	11.8	56.2	20	22	AAH56788	S. aureus groE ope
C 34	11.8	56.2	20	22	AAH56789	S. aureus groE ope
C 35	11.8	56.2	20	24	ABK23143	Human Znaxi cDNA f
C 36	11.8	56.2	20	25	ACC45726	Human HEM SVS mark
C 37	11.8	56.2	21	15	AAQ71069	Primer #1 for ampl
C 38	11.6	55.2	18	18	AAAT91792	Primer DET643 for
C 39	11.6	55.2	19	14	AAQ51831	bcr mRNA ribozyme
C 40	11.6	55.2	20	16	AAQ75163	RAR-beta-1/beta-3
C 41	11.6	55.2	20	18	AAAT66286	Primer 30 for hop
C 42	11.6	55.2	20	20	AAAX59763	Primer P2 used to
C 43	11.6	55.2	20	21	AAZ92612	Primer 4, a murine
C 44	11.6	55.2	20	22	AAF89928	PCR primer used to
C 45	11.6	55.2	20	24	AAAD29426	Human BMPR2 exon1

ALIGNMENTS

RESULT 1
ABX95022/c
ID ABX95022 standard; DNA; 21 BP.

XX AC ABX95022;

XX DT 05-JUN-2003 (first entry)

XX XX Human MAGE-C3 expression pattern anlalysis RT-PCR antisense primer.

XX TRAP; ss; tumour rejection antigen precursor; cytolytic T-cell; CTL;
XX tumour; seminoma; bladder transitional-cell carcinoma; NSCLC; adaptor;
XX head-and-neck squamous-cell carcinoma; breast carcinoma; sarcoma;
XX cutaneous melanoma; nonsmall cell lung cancer; MAGE-C3;
XX human; reverse transcription.

XX OS Homo sapiens.

XX XX US2002176965-A1.

XX XX 28-NOV-2002.

XX XX 01-MAR-2002; 2002US-0085108.

XX XX 09-FEB-2000; 2000US-0501104.

PR 25-APR-1997; 97US-0845528.

PR 24-APR-1998; 98US-0066281.

PR 17-DEC-1999; 99US-0468433.

XX XX (LUCA/) LUCAS S.

FA (BOON/) BOON-FALLEUR T.

```

XX Lucas S, Boon-Falleur T;
PI WPI; 2003-328468/31.
XX
XX Novel isolated nucleic acid encoding tumor rejection antigen precursor
PT MAGE-C3, MAGE-B5, or MAGE-B6, useful as diagnostic probes to determine
PT presence of abnormal e.g., tumor cells expressing MAGE-C1, MAGE-B5 or
PT MAGE-B6
XX
XX Example 12; Page 13; 59pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule which encodes
XX a tumor rejection antigen precursor (TRAP) having an amino acid sequence
XX of a TRAP encoded by a fully defined MAGE-C3, MAGE-B5, or MAGE-B6
XX polynucleotide sequence. Also disclosed is a method which is useful for
XX determining presence of cytolytic T-cells specific for complexes of human
XX leukocyte antigen (HLA) and a peptide derived from the nucleic acid in a
XX cytotoxic T-lymphocyte (CTL)-containing sample. The nucleic acid is
XX useful as a diagnostic probe to determine the presence of abnormal
XX (tumour) cells such as seminoma, bladder transitional-cell carcinoma,
XX head-and-neck squamous-cell carcinoma, breast carcinoma, sarcoma,
XX cutaneous melanoma or non-small cell lung cancer (NSCLC) which express
XX MAGE-C1, MAGE-B5 or MAGE-B6. The nucleic acid is useful for diagnosing a
XX disorder characterised by expression of MAGE-C1, MAGE-B5 or MAGE-B6 TRAPs
XX or tumour rejection antigens (TRAs). The present sequence represents the
XX human MAGE-C3 expression pattern analysis reverse transcription (RT)-PCR
XX antisense primer.
XX
XX Sequence 21 BP; 4 A; 7 C; 4 G; 6 T; 0 other;
XX
Query Match 100.0%; Score 21; DB 25; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTGATTCATGACACGAGG 21
Db 21 AAGCTGATTCATGACACGAGG 1

RESULT 2
AAT04143
ID AAT04143 standard; DNA; 19 BP.
XX
XX AAT04143;
XX
XX 19-APR-1996 (first entry)
XX
XX H.pylori.tox (+/-) vacA gene amplification primer 6.
XX
XX Vacuolating toxin; vaccine; immunisation; therapy; mutant; infection;
XX Helicobacter pylori; PCR; primer; amplification; ss.
XX
XX Synthetic.
XX
XX OS WO9522988-A1.
XX
XX 31-AUG-1995.
XX
XX 23-FEB-1995; 95WO-US02219.
XX
XX 23-FEB-1994; 94US-0200232.
XX
XX (UYVA-) UNIV VANDERBILT.
XX
XX Blaser MJ, Cover TL;
XX
XX WPI; 1995-311383/40.
XX
XX Isolated DNA encoding Helicobacter pylori vacuolating toxin - useful
XX for immunisation against H. pylori infection
XX
XX Examples; Page 31; 64pp; English.

```

```

XX Nucleic acid encoding the Helicobacter pylori vacuolating toxin
XX (AAT01432) and a genetically altered mutant strain of H. pylori which
XX contains a foreign nucleic acid and does not express a functional
XX vacuolating toxin may be used to immunise a subject against H.pylori
XX infection. They may possibly also be used therapeutically.
XX The primers AAT04140-5 were used to amplify and compare fragments of the
XX vacuolating toxin coding sequence vacA from tox(+) and tox(-) strains of
XX H.pylori. The primers AAT04142-3 amplified a 0.6 kb region from the
XX downstream portion of the gene in each of the strains tested.
XX
XX Sequence 19 BP; 5 A; 5 C; 3 G; 6 T; 0 other;
XX
Query Match 61.0%; Score 12.8; DB 16; Length 19;
Best Local Similarity 87.5%; Pred. No. 6e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGCTTGATTCATGAC 16
Db 1 AAGCTTGATTCATGAC 16

RESULT 3
AAB2757/C
ID AAB2757 standard; DNA; 19 BP.
XX
XX AAB2757;
XX
XX 04-DEC-2000 (first entry)
XX
XX cdk3 ribozyme binding site #42.
XX
XX Ribozyme; hairpin; hammerhead; gene therapy; vasotropic;
XX restenosis; ss.
XX
XX Mammalia.
XX
XX WO2000032765-A2.
XX
XX 08-JUN-2000.
XX
XX 06-DEC-1999; 99WO-US28772.
XX
XX 04-DEC-1998; 98US-0110954.
XX
XX (IMMU-) IMMUSOL INC.
XX
XX Tritz R, Welch PJ, Barber JR, Robbins JM;
XX
XX WPI; 2000-412314/35.
XX
XX New hairpin and hammerhead ribozyme for inhibiting restenosis, cleaves
XX RNA encoding a cyclin or cell-cycle dependent kinase other than CDK1,
XX PCNA and Cyclin B1
XX
XX Disclosure; Page 51; 109pp; English.
XX
XX The present invention relates to a hairpin or hammerhead ribozyme,
XX designed to cleave RNA encoding a cyclin or cell-cycle dependent kinase
XX other than cell-cycle dependent kinases CDK1, PCNA and Cyclin B1.
XX Representative examples of ribozyme recognition sites are given in
XX AAB2415 to AAB86787. The ribozyme of the invention is useful for
XX inhibiting restenosis by introduction of the ribozyme into cells.
XX The ribozyme is resistant to endonuclease activity and hence is
XX efficient in restenosis treatment.
XX
XX Sequence 19 BP; 6 A; 5 C; 4 G; 4 T; 0 other;
XX
Query Match 61.0%; Score 12.8; DB 21; Length 19;
Best Local Similarity 87.5%; Pred. No. 6e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CTGATTGATGACACGAGG 20

```

Db 19 CTCATTGATGACGAG 4
||| ||||| |||||
RESULT 4
ID AAA82758/c
AC AAA82758 standard; DNA; 19 BP.
XX
AC AAA82758;
XX
DT 04-DEC-2000 (first entry)
XX
DE cdk3 ribozyme binding site #43.
XX
KW Ribozyme; hairpin; hammerhead; gene therapy; vasotropic;
KW restenosis; ss.
XX
OS Mammalia.
XX
PN WO200032765-A2.
XX
PD 08-JUN-2000.
XX
PF 06-DEC-1999; 99WO-US28772.
XX
PR 04-DEC-1998; 98US-0110954.
XX
PA (IMMU-) IMMUSOL INC.
XX
PI Tritz R, Welch PJ, Barber JR, Robbins JM;
XX
XX WPI; 2000-412314/35.
XX
PT New hairpin and hammerhead ribozyme for inhibiting restenosis, cleaves
PT RNA encoding a cyclin or cell-cycle dependent kinase other than CDK1,
PT PCNA and Cyclin B1
XX
XX Disclosure; Page 51; 109pp; English.
PS
XX
XX The present invention relates to a hairpin or hammerhead ribozyme,
CC designed to cleave RNA encoding a cyclin or cell-cycle dependent kinase
CC other than cell-cycle dependent kinases CDK1, PCNA and Cyclin B1.
CC Representative examples of ribozyme recognition sites are given in
CC AAA82415 to AAA86787. The ribozyme of the invention is useful for
CC inhibiting restenosis by introduction of the ribozyme into cells.
CC The ribozyme is resistant to endonuclease activity and hence is
CC efficient in restenosis treatment.
XX
XX Sequence 19 BP; 4 A; 5 C; 4 G; 6 T; 0 other;
SQ
Query Match 61.0%; Score 12.8; DB 21; Length 19;
Best Local Similarity 87.5%; Pred. No. 6e+03; 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 CTCATTGATGACGAG 20
||| ||||| |||||
Db 16 CTCATTGATGACGAG 1
||| ||||| |||||
RESULT 5
ID AAAH57919/c
AC AAAH57919 standard; DNA; 19 BP.
XX
AC AAAH57919;
XX
DT 10-SEP-2001 (first entry)
XX
DE Cell-cycle dependent kinase cdk3 ribozyme binding site SEQ ID NO:343.
XX
KW Human; ribozyme therapy; hairpin ribozyme; hammerhead ribozyme;
KW recognition site; target; ribozyme binding site; eye disease; vulvular;
KW proliferative disease; skin disease; psoriasis; diabetic retinopathy;
KW cytokine; inflammation; cell-cycle dependent kinase; cyclin; MMP;
matrix metalloproteinase; growth factor; reductase; scarring; cytostatic;
antipsoriatic; dermatological; antiseborrheic; antidiabetic; virucide;
antiskilling; ophthalmological; keratolytic; gene therapy; viral wart;
atopic dermatitis; actinic keratosis; squamous cell carcinoma;
basal cell carcinoma; seborrheic wart; vitreoretinopathy; scar;
sickle cell retinopathy; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200130362-A2.
XX
PD 03-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-US29500.
XX
PR 26-OCT-1999; 99US-0161532.
XX
XX (IMMU-) IMMUSOL INC.
XX
XX Robbins JM, Tritz R;
XX WPI; 2001-300427/31.
XX
XX Treating proliferative skin or eye diseases and scarring, using
PT ribozymes that cleave RNA encoding cytokines involved in inflammation,
PT matrix metalloproteinases, growth factors and cell-cycle dependent
PT kinases -
XX
XX Example 1; Page 96; 408pp; English.
PS
XX The present invention describes a method for treating a proliferative
CC skin or eye disease and scarring. The method involves administering a
CC ribozyme (I) which cleaves RNA encoding a cytokine involved in
CC inflammation, matrix metalloproteinase (MMP), cyclin, cell-cycle
CC dependent kinase, growth factor or a reductase, or administering a
CC nucleic acid molecule (II) comprising a promoter operably linked to a
CC nucleic acid segment encoding (I). (I) can have antipsoriatic,
CC dermatological, cytostatic, antiseborrheic, antidiabetic, antisickling,
CC ophthalmological, vulvular, keratolytic and virucide activities, and
CC cleaves RNA encoding cytokine involved in inflammation. (I) can be used
CC in gene therapy. (I) and (II) are useful for treating proliferative
CC skin diseases such as psoriasis, atopic dermatitis, actinic keratosis,
CC squamous or basal cell carcinoma and viral or seborrheic wart. They can
CC also be used for treating proliferative eye diseases such as diabetic
CC retinopathy, vitreoretinopathy, sickle cell retinopathy, retinopathy of
CC prematurity and retinal detachment, and for treating and preventing
CC scarring such as keloid, adhesion and hypertrophic or hypertrophic burn
CC scar. AAH57577 to AAH52099 represent sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 19 BP; 6 A; 5 C; 4 G; 4 T; 0 other;
SQ
Query Match 61.0%; Score 12.8; DB 22; Length 19;
Best Local Similarity 87.5%; Pred. No. 6e+03; 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 CTCATTGATGACGAG 20
||| ||||| |||||
Db 19 CTCATTGATGACGAG 4
||| ||||| |||||
RESULT 6
ID AAAH57920/c
AC AAAH57920 standard; DNA; 19 BP.
XX
AC AAAH57920;
XX
DT 10-SEP-2001 (first entry)
XX
DE Cell-cycle dependent kinase cdk3 ribozyme binding site SEQ ID NO:344.
XX
KW Human; ribozyme therapy; hairpin ribozyme; hammerhead ribozyme;

KW recognition site; target; ribozyme binding site; eye disease; vulnary;
 KW proliferative disease; skin disease; psoriasis; diabetic retinopathy;
 KW cytokine; inflammation; cell-cycle dependent kinase; cyclin; MMP;
 KW matrix metalloproteinase; growth factor; reductase; scarring; cycostatic;
 KW antipsoriatic; dermatological; antiseborrheic; antidiabetic; virucide;
 KW anticaking; ophthalmological; keratolytic; gene therapy; viral wart;
 KW atopic dermatitis; actinic keratosis; squamous cell carcinoma;
 KW basal cell carcinoma; seborrheic wart; vitreoretinopathy; scar;
 KW sickle cell retinopathy; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200130362-A2.
 XX
 XX 03-MAY-2001.
 XX
 XX 26-OCT-2000; 2000WO-US29500.
 XX
 XX 26-OCT-1999; 99US-0161532.
 XX
 XX (IMMU-) IMMUSOL INC.
 XX
 XX Robbins JM, Tritz R;
 XX
 XX WPI; 2001-300427/31.
 XX
 XX Treating proliferative skin or eye diseases and scarring, using
 PT ribozymes that cleave RNA encoding cytokines involved in inflammation,
 PT matrix metalloproteinases, growth factors and cell-cycle dependent
 PT kinases.
 XX
 PS Example 1; Page 97; 408pp; English.
 XX
 CC The present invention describes a method for treating a proliferative
 CC skin or eye disease and scarring. The method involves administering a
 CC ribozyme (I) which cleaves RNA encoding a cytokine involved in
 CC inflammation, matrix metalloproteinase (MMP), cyclin, cell-cycle
 CC dependent kinase, growth factor or a reductase, or administering a
 CC nucleic acid molecule (II) comprising a promoter operably linked to a
 CC nucleic acid segment encoding (I). (I) can have antipsoriatic,
 CC dermatological, cycostatic, antiseborrheic, antidiabetic, anticaking,
 CC ophthalmological, vulnary, keratolytic and virucide activities, and
 CC cleaves RNA encoding cytokine involved in inflammation. (I) can be used
 CC in gene therapy. (I) and (II) are useful for treating proliferative
 CC skin diseases such as psoriasis, atopic dermatitis, actinic keratosis,
 CC squamous or basal cell carcinoma and viral or seborrheic wart. They can
 CC also be used for treating proliferative eye diseases such as diabetic
 CC retinopathy, vitreoretinopathy, sickle cell retinopathy, retinopathy of
 CC prematurity and retinal detachment, and for treating and preventing
 CC scarring such as keloid, adhesion and hypertrophic or hypertrophic burn
 CC scar. AAH57577 to AAH62099 represent sequences used in the
 CC exemplification of the present invention.
 XX
 XX Sequence 19 BP; 4 A; 5 C; 4 G; 6 T; 0 other;
 SQ
 Query Match 61.0%; Score 12.8; DB 22; Length 19;
 Best Local Similarity 87.5%; Pred. No. 6e+03;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 5 CTGATTGATGACACAGG 20
 ||| ||||| |||||
 Db 16 CTCATTGATGACAGG 1
 RESULT 7
 AAQ95757/c
 ID AAQ95757 standard; DNA; 20 BP.
 XX
 AC AAQ95757;
 XX
 XX 20-FEB-1996 (first entry)
 DT
 XX

DE Primer B (Group 8, set A) for marker D11S910, chromosome 11.
 XX
 KW primer; polymerase chain reaction; PCR; linkage study; locus;
 KW microsatellite marker sequence; automated genotyping; allele;
 KW polymorphism; detection; Homo sapiens; ss.
 OS Synthetic.
 XX
 PN WO9515400-A1.
 XX
 XX 08-JUN-1995.
 XX
 XX 05-DEC-1994; 94WO-US13945.
 XX
 XX 03-DEC-1993; 93US-0160837.
 XX
 XX (UYJO) UNIV JOHNS HOPKINS.
 XX
 XX Levitt RC;
 XX
 XX WPI; 1995-215278/28.
 XX
 XX Kit for automated genotyping contg. pairs of PCR primers - designed
 PT to amplify polymorphic nucleotide repeat sequences, arranged in sets
 PT each with a characteristic fluorescence label, useful e.g. in
 PT detection of disease related genetic rearrangement
 XX
 XX Disclosure; Fig 7H-3; 104pp; English.
 PS
 XX The method aims to provide a collection of highly reproducible
 CC microsatellite marker sequences (MMS) at approx. 10-50 cm intervals
 CC throughout the human genome which can be detectably labelled. The
 CC MMS are polymorphic, simple sequence repeats and can be used in
 CC automated genotyping. esp. fluorescence-based. The primers correspond
 CC to the unique DNA sequence surrounding each marker, and PCR is used to
 CC detect each polymorphism. When the MMS show considerable polymorphism
 CC (ie. a difference in the number of repeats) between individuals, the
 CC markers can be particularly informative. The MMS can be ideal for
 CC linkage studies. Kits comprise at least 4 groups, of at least 3 sets,
 CC each comprising labelled primers for PCR amplification of the DNA.
 CC Group 8 primer pairs are shown in AAQ95735-76. The published size range
 CC of the D11S910 allele is 249-261 bp, and the degree of heterozygosity
 CC in the population is about 71%.
 XX
 XX Sequence 20 BP; 4 A; 8 C; 3 G; 5 T; 0 other;
 SQ
 Query Match 60.0%; Score 12.6; DB 16; Length 20;
 Best Local Similarity 78.9%; Pred. No. 7.7e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 GCCTGATTGATGACACAGG 21
 ||| ||||| |||||
 Db 20 GGCTGACTTATGACAGG 2
 RESULT 8
 AAQ93998
 ID AAQ93998 standard; DNA; 21 BP.
 XX
 AC AAQ93998;
 XX
 XX 08-JUN-1998 (first entry)
 DT
 XX Chicken matrix metalloproteinase-2 gene 5' PCR primer.
 DE
 XX Matrix metalloproteinase-2; MMP-2; chMMP-2; chicken;
 XX Angiogenesis; inhibitor; antagonist; integrin alpha-v beta-3;
 KW vitronectin receptor; rheumatoid arthritis; tumour; metastasis;
 KW diabetic retinopathy; macular degeneration; restenosis; therapy;
 KW PCR; primer; ss.
 XX
 XX Synthetic.
 OS
 OS Gallus sp.

XX WO9745137-A1.

XX PD 04-DEC-1997.

XX PF 30-MAY-1997; 97WO-US09158.

XX PR 31-MAY-1996; 96US-0018733.

XX PR 31-MAY-1996; 96US-0015869.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Brooks P, Cheresh DA;

XX DR WPI; 1998-032334/03.

XX PT Packaging material containing polypeptide antagonist of alphav,
PT beta3 integrin - used for inhibition of angiogenesis, and for
PT treating tumours, inflammation, eye diseases etc.

XX PS Example 4; Page 59; 234pp; English.

XX CC This 5' primer corresponds to nucleotides 809-830 of the chicken
CC matrix metalloproteinase-2 (chMMP-2) mature polypeptide coding
CC region (see AAV03995). It was used with a 3' primer (see AAV03996)
CC to amplify a portion of coding region encoding amino acids
CC 274-637 of chMMP-2. The 5' primer was designed to incorporate a
CC BamHI site allowing directional ligation of the amplified product
CC into pGEX-lambda1 or pGEX-3X expression vectors. Recombinant
CC chMMP-2 (274-637) was expressed as a fusion protein with
CC glutathione-S-transferase in E. coli transformants. The invention
CC relates to the discovery that angiogenesis is mediated by the
CC specific vitronectin receptor alpha-v beta-3, and that inhibition
CC of alpha-v beta-3 function inhibits angiogenesis. Claimed
CC antagonists of alpha-v beta-3 include C-terminal fragments (see
CC AAW41083-94) of human or chicken MMP-2 and fusion proteins.

XX SQ Sequence 21 BP; 5 A; 6 C; 6 G; 4 T; 0 other;

Query Match 60.0%; Score 12.6; DB 19; Length 21;
Best Local Similarity 78.9%; Pred. No. 7.8e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GCCTGATTGATGACCAAGG 21

Db 1 GCCGGATCCATGACCAAGT 19

RESULT 9

AAV12875

ID AAV12875 standard; DNA; 21 BP.

XX AC AAV12875;

XX DT 09-JUN-1998 (first entry)

XX DE 5' PCR primer used to obtain cDNA encoding chicken MMP-2 fragments.

XX KW Metalloproteinase-2; MMP-2; alpha-v-beta-5 antagonist; treatment;
XX vitronectin receptor; inhibition; angiogenesis; tumour growth;
XX restenosis; neovascularisation; PCR primer; ss.

XX OS Synthetic.

XX OS Gallus sp.

XX PN WO9745447-A1.

XX XX 04-DEC-1997.

XX PD 30-MAY-1997; 97WO-US09099.

XX PF 31-MAY-1996; 96US-0018733.

XX PR 31-MAY-1996; 96US-0015869.

XX PA

XX PI (SCRI) SCRIPPS RES INST.

XX PI Brooks P, Cheresh DA, Friedlander M;

XX DR WPI; 1998-041758/04.

XX XX

XX PT Packaging material containing polypeptide antagonist of alphav,
PT beta5 integrin - used for inhibition of angiogenesis, and for
PT treating tumours, inflammation, eye diseases etc.

XX PS

XX PS Example 7; Page 57; 117pp; English.

XX XX

XX CC PCR primers AAV12874-78 were used, together with AAV12873 as the 3'
CC primer, to obtain cDNA encoding chicken matrix metalloproteinase-2 (MMP-2)
CC protein fragments (AAW41234-39). Fragments of the MMP-2 protein act as
CC alpha-v-beta-5 antagonists. Alpha-v-beta-5 is a vitronectin receptor.
CC Inhibitors of alpha-v-beta-5 can inhibit angiogenesis. The specification
CC describes a novel labelled package that contains an inhibitor of
CC angiogenesis i.e. an alpha-v-beta-5 antagonising polypeptide that binds
CC to integrin alpha-v-beta-5 and includes a part of the C-terminal domain
CC of MMP-2. The antagonists are used to inhibit angiogenesis in inflamed
CC tissue, in solid tumours or metastases, and in a wide range of ocular
CC disorders (e.g. diabetic or other forms of retinopathy, neovascular
CC glaucoma, or corneal transplants). They are particularly used to induce
CC regression or to inhibit growth of tumours. The alpha-v-beta-5
CC antagonists can also be used to treat restenosis caused by migration of
CC smooth muscle cells following angioplasty and to reduce blood supply to
CC selected tissues. The antagonists particularly inhibit
CC neovascularisation where this is induced by cytokines, e.g. transforming
CC growth factor alpha, epidermal growth factor or especially vascular
CC endothelial growth factor.

XX SQ Sequence 21 BP; 5 A; 6 C; 6 G; 4 T; 0 other;

Query Match 60.0%; Score 12.6; DB 19; Length 21;
Best Local Similarity 78.9%; Pred. No. 7.8e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GCCTGATTGATGACCAAGG 21

Db 1 GCCGGATCCATGACCAAGT 19

RESULT 10

AAF96809/c

ID AAF96809 standard; DNA; 21 BP.

XX AC AAF96809;

XX DT 06-JUN-2001 (first entry)

XX DE Human gene single nucleotide polymorphism #1570.

XX KW Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
XX polymorphism; vascular disease; coronary artery disease; forensics;
XX myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
XX pulmonary embolism; paternity test; ds.

XX OS Homo sapiens.

XX XX

XX FT Key

XX FT Variation

XX FT replace(11,G)

XX FT /*tag= a

XX FT /standard_name= "single nucleotide polymorphism"

XX PN WO200118250-A2.

XX XX 15-MAR-2001.

XX PD 07-SEP-2000; 2000WO-US24503.

XX PF 10-SEP-1999; 99US-0153357.

PR 26-JUL-2000; 2000US-0220947.
 PR 16-AUG-2000; 2000US-0225724.
 XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PA (MILL-) MILLENNIUM PHARM INC.
 XX Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;
 PI WPI; 2001-226749/23.
 DR Nucleic acids comprising single nucleotide polymorphisms, useful in
 XX applications such as forensics, paternity testing, medicine, genetic
 PT analysis and phenotype correlations to diseases such as diabetes and
 PT atherosclerosis -
 XX Examples; Page 153; 242pp; English.
 PS The present invention provides a method of diagnosing a vascular disease
 XX in an individual, involving determining the sequence at various
 CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4
 CC genes. The sequences at a number of polymorphic sites are also provided
 CC in the specification. In particular, the method can be used in the
 CC diagnosis of atherosclerosis, myocardial infarction, coronary heart
 CC disease, stroke, peripheral vascular diseases, venous thromboembolism
 CC and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
 CC useful in forensics, paternity testing, genetic analysis and phenotype
 CC correlations to diseases. The present sequence is an example of one of
 CC the human gene SNPs shown in the specification.
 XX Sequence 21 BP; 6 A; 9 C; 5 G; 1 T; 0 other;
 SQ Query Match 60.0%; Score 12.6; DB 22; Length 21;
 Best Local Similarity 78.9%; Pred. No. 7.8e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GCCTGATTGATGACACGAGG 21
 DB 20 GCCTGCGTGTGACCTGGG 2

RESULT 11
 AAT50147/c
 ID AAT50147 standard; RNA; 15 BP.
 AC AAT50147;
 AC 07-MAR-1997 (first entry)
 DT Rabbit CETP HH ribozyme target sequence #330.
 DE Hammerhead ribozyme; cholesterol ester transfer protein; mRNA cleavage;
 XX neutral lipid transfer; plasma lipoprotein; atherosclerosis; atherectomy;
 KW reverse cholesterol transport; high density lipoprotein; therapy; CETP;
 KW familial hypercholesterolaemia; dyslipidaemia; hypoalphalipoproteinaemia;
 KW peripheral vascular disease; hyperbetalipoproteinaemia; RCT; inhibitor;
 KW angioplastic restenosis; low density lipoprotein; diabetes; HDL; rabbit;
 KW LDL; ss.
 OS Oryctolagus cuniculus.
 XX WO9620279-A1.
 PN 04-JUL-1996.
 XX 11-DEC-1995; 95WO-US16000.
 PF 23-DEC-1994; 94US-0363240.
 PR (RIBO-) RIBOZYME PHARM INC.
 PA (WARN) WARNER LAMBERT CO.
 XX Bisgaier C, Couture L, McSwiggen J, Pape M, Stinchcomb D;

DR WPI; 1996-321852/32.
 XX New ribozyme(s) for cleaving cholesterol ester transfer protein mRNA
 PT - useful for preventing or treating initial development, progression
 PT or regression of vascular diseases, esp. familial
 PT hypercholesterolaemia
 XX Claim 4; Page 40; 72pp; English.
 PS AAT50138-750359 represent target sequences for the rabbit cholesterol
 CC ester transfer protein (CETP) hammerhead (HH) ribozymes (see
 CC AA50360-750946). CETP is a 74 kD glycoprotein that facilitates neutral
 CC lipid transfer between plasma lipoproteins. The numbering of the targets
 CC refers to the position of the cleavage site in full length CETP. The
 CC ribozyme then binds to 5 nucleotides either side of this site. The
 CC ribozymes are able to cleave mRNA from the gene encoding CETP, thereby
 CC blocking synthesis and/or expression of the mRNA. By inhibiting CETP,
 CC the reverse cholesterol transport (RCT) pathway can be inhibited (or
 CC eliminated) thereby preventing the reduction in size density of the high
 CC density lipoproteins (HDL), prolonging HDL half life, and therefore
 CC increasing HDL levels. The ribozymes can be used to treat conditions
 CC associated with abnormal levels of CETP, specifically atherosclerosis,
 CC familial hypercholesterolaemia, peripheral vascular disease,
 CC dyslipidaemia, hyperbetalipoproteinaemia, hypoalphalipoproteinaemia,
 CC vascular complications of diabetes, transplant, atherectomy and
 CC angioplastic restenosis. By inhibiting CETP, the levels of HDL and low
 CC density lipoproteins (LDL), and the HDL:LDL ratio are favourably altered
 CC (a decrease in LDL levels, and a corresponding increase in HDL levels).
 CC The HH ribozymes can also be used diagnostically to study genetic drift
 CC and mutations in diseased cells, and to detect CETP mRNA. As the HH
 CC ribozymes target specific regions of the CETP gene, they have low
 CC non-specific activity.
 XX Sequence 15 BP; 4 A; 3 C; 4 G; 4 U; 0 other;

QY 5 CTGATTGATGACCA 18
 DB 14 CTGATTGATGCCCA 1

RESULT 12
 AAC64758/c
 ID AAC64758 standard; DNA; 18 BP.
 XX AAC64758;
 AC 28-FEB-2001 (first entry)
 DT RRV PCR primer vMIP-2 SEQ ID NO:167.
 DE Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus;
 KW genome; Kaposi's sarcoma-associated herpesvirus; KSHV; interleukin 6;
 KW IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine;
 KW cytostatic; anti-HIV; gene therapy; infection; Kaposi's sarcoma;
 KW lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy;
 KW splenomegaly; hypergammaglobulinaemia; autoimmune haemolytic anaemia;
 KW PCR primer; ss.
 XX Macaca mulatta rhadinovirus 17577.
 OS WO200028040-A2.
 PN 18-MAY-2000.
 PD 05-NOV-1999; 99WO-US26260.
 PF 06-NOV-1998; 98US-0107507.
 PR 20-NOV-1998; 98US-0109409.
 XX

PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Wong SW, Axthelm MK, Searles RP;
 XX
 DR WPI; 2000-376552/32.
 XX
 PT New rhesus rhadino virus for producing non-human primate model useful
 PT for testing potential treatments and efficacy of the candidate vaccine
 PT for conditions associated with RRV infection -
 XX
 XX Example 13; Page 34; 141pp; English.
 PS
 PS The present invention describes a novel rhesus macaque rhadinovirus
 CC called macaca mulatta rhadinovirus 17577 (RRV). AAC64754 represents the
 CC RRV genome sequence, and AAB53204 represent the proteins
 CC encoded by the genome sequence. The present invention also specifically
 CC claims the individual open reading frame (ORF) nucleotide sequences from
 CC the genome which encode the individual proteins, but these sequences are
 CC not given. A non-human animal infected with RRV can be used for testing
 CC the efficacy of drug in the treatment of condition associated with
 CC infection with RRV such as Kaposi's sarcoma, lymphoproliferative
 CC disorders, B-cell hyperplasia, lymphadenopathy, splenomegaly,
 CC hypergammaglobulinemia or autoimmune haemolytic anaemia, by
 CC administering the drug to a immuno-compromised non-human primate
 CC preferably Rhesus macaque monkey obtained by as a result of infection
 CC by Simian Immunodeficiency Virus (SIV). RRV is useful for producing
 CC non-human primate model for testing potential treatments for conditions
 CC associated with RRV infection. It is also useful for testing the
 CC efficacy of the candidate vaccine against RRV infection or conditions
 CC associated with its infection by administering the vaccine to the
 CC subject capable of infection with RRV, inoculating the subject with RRV
 CC and observing the effect of vaccine. AAC64755 to AAC64765 and AAB53205
 CC to AAB53213 represent sequence used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 18 BP; 4 A; 5 C; 5 G; 4 T; 0 other;
 XX
 Query Match 59.0%; Score 12.4; DB 21; Length 18;
 Best Local Similarity 92.9%; Pred. No. 9.6e+03;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGCCTGATTGATGA 15
 DB 15 AGCCTGATTGACGA 2
 XX
 RESULT 13
 AAV08610
 ID AAV08610 standard; DNA; 19 BP.
 XX
 AC AAV08610;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE Primer ACE/188RT for human ACE gene.
 XX
 KW PCR primer; human; ACE; angiotensin converting enzyme; angiotensinogen;
 KW cardiovascular status; AGT; ATI; type I angiotensin II receptor; stroke;
 KW polymorphic pattern; blood pressure; electrocardiographic profile;
 KW cardiac condition diagnosis; myocardial infarction; atherosclerosis;
 KW hypertension; cardiovascular disease; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX WO9845477-A2.
 PN
 XX 15-OCT-1998.
 PD
 XX 01-APR-1998; 98WO-IB00475.
 PF
 XX 04-APR-1997; 97US-0042930.
 PR
 XX

PA (EURO-) EURONA MEDICAL AB.
 XX
 PI Andersson MK, Lindstroem PHR, Norberg LT;
 XX
 DR WPI; 1998-568361/48.
 XX
 PT Assessing cardiovascular status in humans by polymorphic analysis -
 PT of genes for angiotensin converting enzyme, angiotensinogen and
 PT angiotensin II receptor, used to diagnose predisposition to disease
 PT and to predict effect of therapy
 XX
 XX Example 1; Page 28; 71pp; English.
 PS
 PS This sequence represents a PCR primer for the human ACE (angiotensin
 CC converting enzyme) gene, and can be used in the method of the invention.
 CC The method is for assessing cardiovascular status in humans by
 CC determining the sequence of at least one polymorphic site in the ACE
 CC (angiotensin converting enzyme), AGT (angiotensinogen) and/or ATI (type I
 CC angiotensin II receptor) genes, and comparing the polymorphic pattern
 CC with that in patients with predetermined markers of status. The method is
 CC used to assess blood pressure or electrocardiographic profile, to
 CC diagnose a cardiac condition such as (silent) myocardial infarction (MI),
 CC hypertension, atherosclerosis or stroke. They can also be used to predict
 CC response to treatments with ACE inhibitors, angiotensin II receptor
 CC antagonists, diuretics, alpha- or beta-adrenergic receptor
 CC antagonists, etc. It is also used to identify susceptibility to
 CC cardiovascular disease. Libraries of nucleic acids containing polymorphic
 CC positions in the 3 genes, and libraries of targets corresponding to the
 CC peptides from the genes are used to screen for cardiovascular agents. The
 CC nucleic acids contained in the library can be used as source of
 CC probes.
 XX
 SQ Sequence 19 BP; 3 A; 3 C; 8 G; 5 T; 0 other;
 XX
 Query Match 59.0%; Score 12.4; DB 19; Length 19;
 Best Local Similarity 92.9%; Pred. No. 9.7e+03;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGCCTGATTGATGA 15
 DB 4 AGCCTGGTGTGATGA 17
 XX
 RESULT 14
 AAC61238
 ID AAC61238 standard; DNA; 19 BP.
 XX
 AC AAC61238;
 XX
 DT 30-JAN-2001 (first entry)
 XX
 DE Human ACE, AGT and ATI genes polymorphisms PCR primer SEQ ID NO: 38.
 XX
 KW Human; genetic polymorphism; disease diagnosis; treatment; cancer;
 KW cardiovascular system; nervous system; glaucoma; PCR primer; ss.
 XX
 OS Homo sapiens.
 OS
 XX WO2000056922-A2.
 PN
 XX 28-SEP-2000.
 PD
 XX 23-MAR-2000; 2000WO-GB01102.
 PF
 XX 23-MAR-1999; 99US-0126046.
 PR
 XX 23-MAR-1999; 99WO-IB00497.
 PR
 XX 24-MAR-1999; 99US-0126243.
 PR
 XX 23-DEC-1999; 99US-0471890.
 XX
 XX (GEMI-) GEMINI GENOMICS AB.
 PA
 XX Lindstrom PHR, Norberg LT, Jonsson L, Olaisson E, Sanders R;
 PI

DR WPI; 2000-638268/61..

XX Assessing disease status in individual by determining sequence(s) at

PT one or more polymorphic positions within the human genes encoding the

PT protein(s) involved in physiological pathway associated with treatment

PT regime -

XX Example 1; Page 57; 141pp; English.

PS The present invention is related to methods for determining the

XX polymorphic pattern of an individual and using the results to determine

CC their risk of a number of diseases, including cancer, cardiovascular

CC diseases, glaucoma and nervous system disorders such as depression and

CC neurodegenerative diseases. In addition, the methods can be used to

CC determine the effects of different types of treatment for individuals,

CC and thus enables appropriate therapies to be prescribed. The PCR primers

CC shown in sequences AAC61201-C61371 were all used to demonstrate the

CC methods of the invention.

XX

SQ Sequence 19 BP; 3 A; 3 C; 8 G; 5 T; 0 other;

Query Match 59.0%; Score 12.4; DB 21; Length 19;

Best Local Similarity 92.9%; Pred. No. 9.7e+03;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGCCTGATTGATGA 15

Db 4 AGCCTGGTTGATGA 17

||||| |||||

RESULT 15

AAA38238

ID AAA38238 standard; DNA; 19 BP.

XX AC

XX AAA38238;

XX 21-AUG-2000 (first entry)

XX Human angiotensin-converting enzyme (ACE) PCR primer, SEQ ID NO:38.

XX

XX Angiotensin-converting enzyme gene; ACE; polymorphism;

KW polymorphic marker; cardiovascular disease; myocardial infarction;

KW unstable angina; hypertension; atherosclerosis; stroke; prognosis;

KW drug screening; treatment outcome; human; PCR primer; ss.

XX

OS Homo sapiens.

XX

XX WC200022166-A2.

PN

PD 20-APR-2000.

XX

XX 13-OCT-1999; 99WO-IB01678.

XX

XX 14-OCT-1998; 98US-0104286.

PR 14-OCT-1998; 98US-0104302.

XX

XX (EURO-) EURONA MEDICAL AB.

PA

XX Norberg LT, Andersson MK, Lindstrom PHR, Jonsson L;

PI

XX WPI; 2000-318010/27.

DR

XX Assessing cardiovascular status in humans involves comparing test

PT polymorphic pattern comprising polymorphic positions within genes

PT encoding specific proteins, with reference polymorphic pattern -

XX

XX Example 1; Page 49; 126pp; English.

XX The invention relates to a novel method of assessing the cardiovascular

CC status in an individual and to newly identified polymorphisms in the

CC genes encoding angiotensin-converting enzyme (ACE), angiotensin II

CC receptor type 1 (AT1) and type 2 (AT2), angiotensinogen (AGT), renin,

CC aldosterone synthase, endothelin receptor type A and beta-adrenergic

CC receptors 1 and 2. The method comprises determining the sequence at one

CC or more polymorphic positions within these genes, and comparing the

CC pattern of polymorphisms from the individual with a reference

CC polymorphic pattern obtained from a population of individuals exhibiting

CC a predetermined cardiovascular disease status. The polymorphic markers

CC are useful for determining the predisposition of an individual to

CC cardiovascular disorders such as myocardial infarction, unstable angina,

CC hypertension, atherosclerosis and stroke. They are also useful for

CC predicting the likely cardiovascular status of a patient given a

CC treatment regimen comprising administration of cardiovascular drugs

CC (e.g., ACE inhibitors, beta-adrenergic receptor antagonists (beta-blockers) or calcium channel blockers). One or more polymorphic markers

CC provides a basis for predicting the outcome of a treatment regimen.

CC Fragments of the genes comprising a polymorphic site may be used as

CC primers and probes for detecting genetic polymorphisms or in molecular

CC library arrays for high throughput screening. The genes, and the

CC proteins they encode are useful in the screening of potential

CC cardiovascular drugs. Determination of an individual's polymorphic

CC pattern reduces or eliminates trial and error in selecting a treatment

CC for a particular individual cardiovascular patient. It also provides the

CC ability to eliminate patients from clinical trials who are predicted to

CC be non-responsive, or at a risk for an adverse response, to a particular

CC treatment regimen. Adverse results in an early trial can be evaluated to

CC identify polymorphic patterns so that the adverse results can be

CC correlated with a sub-population of the test population, permitting

CC exclusion of such sub-populations from the treatment group. Beneficial

CC drugs can be approved for use in the appropriate population, thereby

CC decreasing the number of patients required for a clinical trial, which

CC in turn decreases the duration and cost of such trials. Sequences

CC AAA38201-A38239 represent PCR primers used in an exemplification of

CC the invention to amplify short fragments of the human ACE gene

CC (AAA38328- AAA38330) for sequence determination.

XX

SQ Sequence 19 BP; 3 A; 3 C; 8 G; 5 T; 0 other;

Query Match 59.0%; Score 12.4; DB 21; Length 19;

Best Local Similarity 92.9%; Pred. No. 9.7e+03;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGCCTGATTGATGA 15

Db 4 AGCCTGGTTGATGA 17

||||| |||||

Search completed: February 20, 2004, 00:55:58

Job time : 172 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:16:19 ; Search time 1597.5 Seconds
(without alignments)
537.779 Million cell updates/sec

Title: US-10-085-108-21_COPY_711_731

Perfect score: 21

Sequence: 1 AAGCCTGATTGATGACCAAGG 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 637448

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	12.8	61.0	19	6	AX129125	Sequence
C 2	12.8	61.0	19	6	AX129126	Sequence
C 3	12.8	61.0	19	6	189405	Sequence 16
C 4	12.6	60.0	21	6	AR269340	Sequence
C 5	12.6	60.0	21	6	AX096396	Sequence
C 6	12.4	59.0	19	6	AR137291	Sequence
C 7	12.4	59.0	19	6	AX037413	Sequence
C 8	12.4	59.0	19	6	BD075165	Methods f
C 9	12.4	59.0	20	6	AR233464	Sequence
C 10	12.2	58.1	17	6	AX732389	Sequence
C 11	12.2	58.1	17	6	AX735038	Sequence
C 12	12.2	58.1	20	6	AR225911	Sequence
C 13	12	57.1	20	6	AX134104	Sequence
C 14	12	57.1	20	6	AX189713	Sequence
C 15	12	57.1	21	6	AX53656	Sequence
C 16	12	57.1	21	6	BD142983	Method of
C 17	11.8	56.2	17	6	AX735886	Sequence
C 18	11.8	56.2	18	6	AX133050	Sequence
C 19	11.8	56.2	20	6	AX148051	Sequence
C 20	11.8	56.2	20	6	AX149234	Sequence
C 21	11.8	56.2	20	6	AX149235	Sequence
C 22	11.8	56.2	21	4	DOG21302	Sequence
C 23	11.8	56.2	21	6	AR098872	Sequence
C 24	11.8	56.2	21	6	AR275185	Sequence
C 25	11.8	56.2	21	6	I79712	Sequence 7
C 26	11.6	55.2	19	6	A87308	Sequence 33
C 27	11.6	55.2	19	6	BD057386	Protein c
C 28	11.6	55.2	19	6	I83633	Sequence 7
C 29	11.6	55.2	20	6	A95470	Sequence 20
C 30	11.6	55.2	20	6	AR072824	Sequence
C 31	11.6	55.2	20	6	AR177533	Sequence
C 32	11.6	55.2	20	6	AX135802	Sequence
C 33	11.6	55.2	20	6	AX375541	Sequence
C 34	11.6	55.2	21	6	AR100947	Sequence
C 35	11.6	55.2	21	6	AR240235	Sequence
C 36	11.4	54.3	15	6	AX635907	Sequence
C 37	11.4	54.3	15	6	AX636096	Sequence
C 38	11.4	54.3	15	6	I61477	Sequence 31
C 39	11.4	54.3	15	6	I61666	Sequence 22
C 40	11.4	54.3	17	6	AX733387	Sequence
C 41	11.4	54.3	20	6	AX188466	Sequence
C 42	11.4	54.3	20	6	E32494	Sequence
C 43	11.4	54.3	21	6	AX675050	Sequence
C 44	11.4	54.3	21	6	BD088558	Sequence
C 45	11.4	54.3	21	12	AB069455	Synthetic

ALIGNMENTS

RESULT 1	AX129125/c	AX129125	19 bp	DNA	linear	PAT 15-MAY-2001
LOCUS	Sequence 343 from Patent WO0130362.					
DEFINITION	AX129125					
ACCESSION	AX129125					
VERSION	AX129125.1	GI:14135430				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.					
TITLE	Robbins, J.M. and Tritz, R.					
	Ribozyme therapy for the treatment of proliferative skin and eye diseases					

DEFINITION Methods for assessing cardiovascular status and compositions for use thereof.

RESULT 6
AR137291
LOCUS AR137291 19 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 38 from patent US 6197505.
ACCESSION AR137291
VERSION AR137291.1 GI:14478800
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 19)
AUTHORS Norberg,L.T., Andersson,M.K., Lindstrom,P.H., and Lindstrom,P.H., Rutger.
TITLE Methods for assessing cardiovascular status and compositions for use thereof
JOURNAL Patent: US 6197505-A 38 06-MAR-2001;
FEATURES Location/Qualifiers
source 1..19
BASE COUNT 3 a 3 c 8 g 5 t
ORIGIN

Query Match 59.0%; Score 12.4; DB 6; Length 19;
Best Local Similarity 92.9%; Pred. No. 3.7e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGCCTGATTGATGA 15
|||||
Db 4 AGCCTGGTTGATGA 17

RESULT 7
AX037413
LOCUS AX037413 19 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 38 from Patent WO0056922.
ACCESSION AX037413
VERSION AX037413.1 GI:11226838
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Norberg,L.T., Olaisson,E., Jonsson,L., Lindstrom,P.H. and Sanders,R.
TITLE Genetic polymorphism and polymorphic pattern for assessing disease status, and compositions for use thereof
JOURNAL Patent: WO 0056922-A 38 28-SEP-2000;
NORBERG LEIF TORBJORN (SE) ; OLAISSON ERIK (SE) ; JONSSON LENA (SE) ; GEMINI GENOMICS AB (SE) ; LINDSTROM PER HARRY RUTGER (SE) ; SANDERS RHIANNON (SE)
FEATURES Location/Qualifiers
source 1..19
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide primer"

BASE COUNT 3 a 3 c 8 g 5 t
ORIGIN
Query Match 59.0%; Score 12.4; DB 6; Length 19;
Best Local Similarity 92.9%; Pred. No. 3.7e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGCCTGATTGATGA 15
|||||
Db 4 AGCCTGGTTGATGA 17

RESULT 8
BD075165
LOCUS BD075165 19 bp DNA linear PAT 27-AUG-2002

DEFINITION Methods for assessing cardiovascular status and compositions for use thereof.

ACCESSION BD075165
VERSION BD075165.1 GI:22620768
KEYWORDS JP 2001519660-A/38.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 19)
AUTHORS Norberg,L.T., Andersson,M.K. and Lindstrom,P.H.R.
TITLE Methods for assessing cardiovascular status and compositions for use thereof
JOURNAL Patent: JP 2001519660-A 38 23-OCT-2001;
COMMENT OS Artificial Sequence
PN JP 2001519660-A/38
PD 23-OCT-2001
PF 01-APR-1998 JP 1998542530
PR 04-APR-1997 US 60/042930
PI LEIF TORBJORN NORBERG,MARIA KRISTINA ANDERSSON,PER HARRY PI RUTGER LINDSTROM
PC C1201/68,C07K14/72,C07K14/575,C12N9/48
CC Description of Artificial Sequence: PCR PRIMER FH Key
FEATURES Location/Qualifiers
FT source 1..19
FT /organism='Artificial Sequence'.
source Location/Qualifiers
1..19
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

BASE COUNT 3 a 3 c 8 g 5 t
ORIGIN

Query Match 59.0%; Score 12.4; DB 6; Length 19;
Best Local Similarity 92.9%; Pred. No. 3.7e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGCCTGATTGATGA 15
|||||
Db 4 AGCCTGGTTGATGA 17

RESULT 9
AR233464
LOCUS AR233464 20 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 93 from patent US 6458532.
ACCESSION AR233464
VERSION AR233464.1 GI:27276055
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 20)
AUTHORS Detera-Wadleigh,S.D., Yoshikawa,T., Sanders,A.R. and Esterling,L.E.
TITLE Polynucleotides encoding IMP 18p myo-inositol monophosphatase and methods of detecting said polynucleotides
JOURNAL Patent: US 6458532-A 93 01-OCT-2002;
FEATURES Location/Qualifiers
source 1..20
/organism="unknown"

BASE COUNT 5 a 2 c 9 g 4 t
ORIGIN
Query Match 59.0%; Score 12.4; DB 6; Length 20;
Best Local Similarity 92.9%; Pred. No. 3.7e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GATTGATGACCAGG 20
|||||
Db 2 GATTGAGGACCAGG 15

```

RESULT 10
AX732389/c
LOCUS AX732389 17 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 4023 from Patent WO03025175.
ACCESSION AX732389
VERSION AX732389.1 GI:30511732
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
FEATURES
source Telerman,A., Amson,R. and Tuijnder,M.
Patent: WO 03025175-A 4023 27-MAR-2003;
Molecular Engines Laboratories (FR)
Location/Qualifiers
1..17
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 6 a 4 c 4 g 3 t
ORIGIN
Query Match 58.1%; Score 12.2; DB 6; Length 17;
Best Local Similarity 82.4%; Pred. No. 4.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACC 17
Db 17 ATGCCTGATTGCTGATC 1

RESULT 11
AX735038/c
LOCUS AX735038 17 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 628 from Patent WO03025177.
ACCESSION AX735038
VERSION AX735038.1 GI:30514315
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
FEATURES
source Telerman,A., Amson,R. and Tuijnder,M.
Sequences involved in phenomena of tumour suppression, tumour
reversion, apoptosis and/or resistance to viruses and the use
thereof as medicaments
Patent: WO 03025177-A 628 27-MAR-2003;
Molecular Engines Laboratories (FR)
Location/Qualifiers
1..17
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 6 a 4 c 4 g 3 t
ORIGIN
Query Match 58.1%; Score 12.2; DB 6; Length 17;
Best Local Similarity 82.4%; Pred. No. 4.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACC 17
Db 17 ATGCCTGATTGCTGATC 1

RESULT 12
AR225911/c
LOCUS AR225911 20 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 15 from Patent WO0148240.
ACCESSION AX189713
VERSION AX189713.1 GI:15143089
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
FEATURES
source Telerman,A., Amson,R. and Tuijnder,M.
Sequences involved in phenomena of tumour suppression, tumour
reversion, apoptosis and/or resistance to viruses and the use
thereof as medicaments
Patent: WO 03025177-A 628 27-MAR-2003;
Molecular Engines Laboratories (FR)
Location/Qualifiers
1..17
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 6 a 4 c 4 g 3 t
ORIGIN
Query Match 58.1%; Score 12.2; DB 6; Length 17;
Best Local Similarity 82.4%; Pred. No. 4.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACC 17
Db 17 ATGCCTGATTGCTGATC 1

```

```

DEFINITION Sequence 61 from patent US 6444464.
ACCESSION AR225911
VERSION AR225911.1 GI:27264065
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Wyatt,J.
TITLE Antisense modulation of E2F transcription factor 2 expression
JOURNAL Patent: US 6444464-A 61 03-SEP-2002;
FEATURES
source Location/Qualifiers
1..20
/organism="unknown"
BASE COUNT 2 a 6 c 6 g 6 t
ORIGIN
Query Match 58.1%; Score 12.2; DB 6; Length 20;
Best Local Similarity 82.4%; Pred. No. 4.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCTGATTGATGACGAGG 20
Db 18 CCTGACTGAGCACAGG 2

RESULT 13
AX134104/c
LOCUS AX134104 20 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 15 from Patent EP1113081.
ACCESSION AX134104
VERSION AX134104.1 GI:14270868
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
FEATURES
source Charrier-Harlin,M.C., Amouyel,P. and Lambert,J.C.
Patent: EP 1113081-A 15 04-JUL-2001;
Institute Pasteur de Lille (FR); Institut National de la Sante et
de la Recherche Medicale (INSERM) (FR)
Location/Qualifiers
1..20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 3 a 6 c 4 g 7 t
ORIGIN
Query Match 57.1%; Score 12; DB 6; Length 20;
Best Local Similarity 75.0%; Pred. No. 5.9e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AGCCTGATTGATGACGAGG 21
Db 20 AGCCAGATTTCACAGCAGG 1

RESULT 14
AX189713/c
LOCUS AX189713 20 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 15 from Patent WO0148240.
ACCESSION AX189713
VERSION AX189713.1 GI:15143089
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
FEATURES
source Telerman,A., Amson,R. and Tuijnder,M.
Sequences involved in phenomena of tumour suppression, tumour
reversion, apoptosis and/or resistance to viruses and the use
thereof as medicaments
Patent: WO 03025177-A 628 27-MAR-2003;
Molecular Engines Laboratories (FR)
Location/Qualifiers
1..17
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 6 a 4 c 4 g 3 t
ORIGIN
Query Match 58.1%; Score 12.2; DB 6; Length 17;
Best Local Similarity 82.4%; Pred. No. 4.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACC 17
Db 17 ATGCCTGATTGCTGATC 1

```

AUTHORS Chartier-Harlin,M.C., Amouyel,P., Lambert,J.C. and Araria,L.
TITLE Implication of a known gene named cp2/lstf-lbp-1 in Alzheimer's
disease
JOURNAL Patent: WO 0148240-A 15 05-JUL-2001;
INSTITUT PASTEUR DE LILLE (FR) ; INSTITUT NATIONAL DE LA SANTE ET
DE LA RECHERCHE MEDICALE (INSERM) (FR)

FEATURES
source
1. .20
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 3 a 6 c 4 g 7 t
ORIGIN

Query Match 57.1%; Score 12; DB 6; Length 20;
Best Local Similarity 75.0%; Pred. No. 5.9e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AGCGTGATGATGACACGAGG 21
|||||
Db 20 AGCCAGATTTACACGACGAGG 1

RESULT 15
AX553656/c
LOCUS AX553656 21 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 60 from Patent WO02074946.
ACCESSION AX553656
VERSION AX553656.1 GI:25897654
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Serup,P., Heimberg,H. and Gradwohl,G.
TITLE Method for generating insulin-secreting cells suitable for
transplantation
JOURNAL Patent: WO 02074946-A 60 26-SEP-2002;
NOVO NORDISK A/S (DK)
FEATURES Location/Qualifiers
source
1. .21
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 6 a 6 c 4 g 5 t
ORIGIN

Query Match 57.1%; Score 12; DB 6; Length 21;
Best Local Similarity 75.0%; Pred. No. 5.9e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGCGTGATGATGACACGAGG 20
|||||
Db 21 AGCGCTGTTTGTGACCAAG 2

Search completed: February 20, 2004, 01:49:33
Job time : 1599.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:22:43 ; Search time 2275.5 Seconds
(without alignments)
306.060 Million cell updates/sec

Title: US-10-085-108-21_COPY_711_731

Perfect score: 21

Sequence: 1 AAGCTGATTGATGACCAAGG 21

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 33363688 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:*

```
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*old.*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
```

```
44: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	100.0	425	1	PCT-US01-00663-4923 Sequence 4923, Ap
C 2	21	100.0	425	2	PCT-US01-00663-4923 Sequence 4923, Ap
C 3	21	100.0	425	36	US-09-864-761-4696 Sequence 4696, Ap
C 4	21	100.0	425	47	US-10-182-993-4749 Sequence 4749, Ap

RESULT 3
US-09-864-761-4596/c
; Sequence 4596, Application US/09864761
; GENERAL INFORMATION:

```

PCT-US01-00663-4923/C
; Sequence 4923, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408

```

```

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4636
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL023279.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.3
US-09-864-761-4696
Query Match 100.0%; Score 21; DB 36; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AAGCCTGATTGATGACCGG 21
Db 150 AAGCCTGATTGATGACCGG 130

```

```

RESULT 4
US-10-182-993-4749/c
; Sequence 4749, Application US/10182993
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 2
; CURRENT APPLICATION NUMBER: US/10/182,993
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37811
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 4749
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL023279.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
US-10-182-993-4749
Query Match 100.0%; Score 21; DB 47; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AAGCCTGATTGATGACCGG 21
Db 150 AAGCCTGATTGATGACCGG 130

```

```

RESULT 5
US-10-182-995-4696/c
; Sequence 4696, Application US/10182995
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 1
; CURRENT APPLICATION NUMBER: US/10/182,995
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359

```

;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 29119
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4696
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HEART, SIGNAL = 6
;; FILE REFERENCE: PB 0004 WO 9
;; CURRENT APPLICATION NUMBER: US/10/182,998
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 26941
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4815
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HELA, SIGNAL = 7.1
;; FILE REFERENCE: PB 0004 WO 6
;; CURRENT APPLICATION NUMBER: US/10/203,134
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)

Query Match 100.0%; Score 21; DB 47; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21
|||
Db 150 AAGCTGATTGATGACCAGG 130

RESULT 6
US-10-182-997-4815/C
;; Sequence 4815, Application US/10182997
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 10
;; CURRENT APPLICATION NUMBER: US/10/182,997
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 26941
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4815
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HELA
;; FILE REFERENCE: PB 0004 WO 10
;; CURRENT APPLICATION NUMBER: US/10/182,997
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 26941
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4815
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HELA, SIGNAL = 7.1
;; FILE REFERENCE: PB 0004 WO 6
;; CURRENT APPLICATION NUMBER: US/10/203,134
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)

Query Match 100.0%; Score 21; DB 47; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21
|||
Db 150 AAGCTGATTGATGACCAGG 130

RESULT 7
US-10-182-998-4657/C
;; Sequence 4657, Application US/10182998
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 9
;; CURRENT APPLICATION NUMBER: US/10/182,998
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; NUMBER OF SEQ ID NOS: 15009
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4657
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HELA
;; FILE REFERENCE: PB 0004 WO 9
;; CURRENT APPLICATION NUMBER: US/10/182,998
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)

Query Match 100.0%; Score 21; DB 47; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21
|||
Db 150 AAGCTGATTGATGACCAGG 130

RESULT 8
US-10-203-134-4839/C
;; Sequence 4839, Application US/10203134
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 6
;; CURRENT APPLICATION NUMBER: US/10/203,134
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; NUMBER OF SEQ ID NOS: 15009
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4657
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HELA
;; FILE REFERENCE: PB 0004 WO 6
;; CURRENT APPLICATION NUMBER: US/10/203,134
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)

Query Match 100.0%; Score 21; DB 47; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21
|||
Db 150 AAGCTGATTGATGACCAGG 130

;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38628
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4839
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.2
US-10-203-134-4839

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCAGG 21
DB 150 AAGCCTGATTGATGACCAGG 130

RESULT 9
US-10-203-135-4857/c
;; Sequence 4857, Application US/10203135
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 5
;; CURRENT APPLICATION NUMBER: US/10/203,135
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 37012
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4857
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
US-10-203-135-4857

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCAGG 21
DB 150 AAGCCTGATTGATGACCAGG 130

DB 150 AAGCCTGATTGATGACCAGG 130
RESULT 10
US-10-203-136-4917/c
;; Sequence 4917, Application US/10203136
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 3
;; CURRENT APPLICATION NUMBER: US/10/203,136
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38578
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4917
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.9
US-10-203-136-4917

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCAGG 21
DB 150 AAGCCTGATTGATGACCAGG 130

RESULT 11
US-10-203-137-4923/c
;; Sequence 4923, Application US/10203137
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 7
;; CURRENT APPLICATION NUMBER: US/10/203,137
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
US-10-203-137-4923

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38937
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4923
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;;
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.3
US-10-203-137-4923

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21
|||||
Db 150 AAGCTGATTGATGACCAGG 130

RESULT 12
US-10-203-138-4770/c
;; Sequence 4770, Application US/10203138
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 8
;; CURRENT APPLICATION NUMBER: US/10/203,138
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; NUMBER OF SEQ ID NOS: 15438
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4770
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;;
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4
US-10-203-138-4770

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21
|||||
Db 150 AAGCTGATTGATGACCAGG 130

RESULT 14
US-10-203-139-4922/c
;; Sequence 4922, Application US/10203139
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 4
;; CURRENT APPLICATION NUMBER: US/10/203,139
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; NUMBER OF SEQ ID NOS: 15438
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4770
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;;
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4
US-10-203-139-4770

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21

Db 150 AAGCTGATTGATGACCAGG 130
|||||

RESULT 13
US-10-203-138A-4770/c
;; Sequence 4770, Application US/10203138A
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 8
;; CURRENT APPLICATION NUMBER: US/10/203,138A
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; NUMBER OF SEQ ID NOS: 15438
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4770
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;;
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4
US-10-203-138A-4770

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTGATTGATGACCAGG 21
|||||
Db 150 AAGCTGATTGATGACCAGG 130

RESULT 14
US-10-203-139-4922/c
;; Sequence 4922, Application US/10203139
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 4
;; CURRENT APPLICATION NUMBER: US/10/203,139
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)

Db 150 AAGCCTGATTGATGACCGGG 130

Search completed: February 20, 2004, 00:32:19
Job time : 2278.5 secs

;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 37156
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 4922
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.7
US-10-203-139-4922

Query Match 100.0%; Score 21; DB 48; Length 425;
Best Local Similarity 100.0%; Pred.No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCGGG 21
|||||
Db 150 AAGCCTGATTGATGACCGGG 130

RESULT 15
US-60-236-359-5855/c
;; Sequence 5855, Application US/60236359
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: MDHMOF-4P
;; CURRENT APPLICATION NUMBER: US/60/236,359
;; CURRENT FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; NUMBER OF SEQ ID NOS: 21709
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 5855
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.2
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.9
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.7
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.3
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.3
;; OTHER INFORMATION: EST HUMAN HIT: BE276397.1, EVALUATE 8.00e-11
;; OTHER INFORMATION: NT HIT: q1 4885474, EVALUATE 2.00e-78
;; OTHER INFORMATION: SWISSPROT HIT: P43363, EVALUATE 1.00e-42
US-60-236-359-5855

Query Match 100.0%; Score 21; DB 78; Length 425;
Best Local Similarity 100.0%; Pred.No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCGGG 21
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 22:22:13 ; Search time 194 seconds
(without alignments)
398.744 Million cell updates/sec

Title: US-10-085-108-21_COPY_711_731
Perfect score: 21
Sequence: 1 AAGCTGATGATGACACGGG 21

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2449703 seqs, 184181367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	100.0	425	9	US-09-864-761-4696
C 2	21	100.0	921	13	US-10-027-632-100997
C 3	21	100.0	921	14	US-10-027-632-100997
C 4	21	100.0	1041	14	US-10-085-108-21
C 5	17.4	82.9	1022	13	US-10-027-632-258115
C 6	17.4	82.9	1022	14	US-10-027-632-258115
C 7	16.8	80.0	504	10	US-09-728-445-868
C 8	16.8	80.0	12591	15	US-10-213-878-6
C 9	16.8	80.0	12591	15	US-10-214-059-6
C 10	16.2	77.1	456	9	US-09-864-761-11491
C 11	16.2	77.1	481	11	US-09-918-995-283
C 12	16.2	77.1	512	13	US-10-027-632-37978
C 13	16.2	77.1	512	13	US-10-027-632-76583
C 14	16.2	77.1	512	14	US-10-027-632-37978
C 15	16.2	77.1	512	14	US-10-027-632-76583

16	16.2	77.1	619	13	US-10-027-632-32980
17	16.2	77.1	619	14	US-10-027-632-32980
18	16.2	77.1	741	13	US-10-027-632-151775
19	16.2	77.1	741	14	US-10-027-632-151775
20	16.2	77.1	1657	13	US-10-027-632-255152
21	16.2	77.1	1657	14	US-10-027-632-255152
22	16.2	77.1	1920	13	US-10-181-157-4
23	16.2	77.1	4912	11	US-09-843-250-12
24	16.2	77.1	6779	11	US-09-843-250-13
25	16.2	77.1	7325	13	US-10-240-965-72
26	15.8	75.2	441	12	US-10-369-493-40614
27	15.8	75.2	470	9	US-09-864-761-5173
28	15.8	75.2	707	13	US-10-027-632-12506
29	15.8	75.2	707	14	US-10-027-632-12506
30	15.8	75.2	809	13	US-10-027-632-153398
31	15.8	75.2	809	14	US-10-027-632-153398
32	15.8	75.2	1343	10	US-09-887-576-481
33	15.8	75.2	1344	10	US-09-887-576-576
34	15.8	75.2	3457	12	US-10-108-260A-1224
35	15.4	73.3	312	10	US-09-738-626-1944
36	15.4	73.3	320	9	US-09-764-887-17
37	15.4	73.3	320	15	US-10-073-961-17
38	15.4	73.3	1185	12	US-10-369-493-43880
39	15.4	73.3	2377	12	US-10-369-493-27069
40	15.4	73.3	2753	9	US-09-764-887-323
41	15.4	73.3	2753	9	US-09-764-887-324
42	15.4	73.3	2753	15	US-10-073-961-323
43	15.4	73.3	2753	15	US-10-073-961-324
44	15.4	73.3	4780	9	US-09-886-241-1
45	15.4	73.3	3309400	10	US-09-738-626-1

ALIGNMENTS

RESULT 1

US-09-864-761-4696/c
; Sequence 4696, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 4696
;; LENGTH: 425
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL023279.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.2
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.9
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.7
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.3
US-09-864-761-4696

Query Match 100.0%; Score 21; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCCAGG 21
Db 150 AAGCCTGATTGATGACCCAGG 130

RESULT 2
US-10-027-632-100997/c
; Sequence 100997, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100997
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-100997

Query Match 100.0%; Score 21; DB 14; Length 921;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCCAGG 21
Db 156 AAGCCTGATTGATGACCCAGG 136

RESULT 4
US-10-085-108-21
; Sequence 21, Application US/10085108
; Publication No. US2002017686SAI
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING
; FOR
; TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-(
; MAGE-B FAMILIES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS

Query Match 100.0%; Score 21; DB 13; Length 921;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCCAGG 21
Db 156 AAGCCTGATTGATGACCCAGG 136

RESULT 3
US-10-027-632-100997/c
; Sequence 100997, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100997
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-100997

SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/085.108
FILING DATE: 01-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/501.104
FILING DATE: 09-Feb-2000
APPLICATION NUMBER: 09/468.433
FILING DATE: December 17, 1999
APPLICATION NUMBER: 09/066.281
FILING DATE: April 24, 1998
APPLICATION NUMBER: 08/845.528
FILING DATE: April 25, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mary Anne Schofield
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/MAS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3100
TELEFAX: (212) 318-3400
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1041 base pairs
TYPE: nucleic acid
STRANDEDNESS: SINGLE
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-085-108-21

```

Query Match      100.0%; Score 21; DB 14; Length 1041;
Best Local Similarity 100.0%; P: 0.61;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATCACCAGG 21
    |||||
Db 711 AAGCCTGATTGATCACCAGG 731

```

```

RESULT 5
US-10-027-632-258115/c
; Sequence 258115, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258115
; LENGTH: 1022
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258115

```

Query Match 82.9%; Score 17.4; DB 13; Length 1022;

```

Best Local Similarity   94.7%;   Pred: No. 47;
Matches    18; Conservative    0; Mismatches      1; Indels      0; Gaps      0;

QY          3 GCCTGATTGATCACCAGG 21
            |||||
Db          969 GCCTGATTCATGACCAGG 951

RESULT 6
US-10-027-632-258115/c
; Sequence 258115, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258115
; LENGTH: 1022
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258115

```

```

Query Match      82.9%; Score 17.4; DB 14; Length 1022;
Best Local Similarity 94.7%; Pred.No. 47;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      3  GCCTGATTGATGACCAAGG 21
          |||||
Db      969  GCCTGATTCAIGACCAAGG 951

```

RESULT 7
US-09-728-445-868
; Sequence 868, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. US20020102543A1el Mutated Mammalian Cells and
; TITLE OF INVENTION: Animals
; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 868
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-728-445-868

Query Match 80.0%; Score 16.8; DB 10; Length 504;

Best Local Similarity 90.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCCTGATTGATGACCCAGG 21
Db 461 AGCAGGATTGATGACCCAGG 480

RESULT 8

US-10-213-878-6
; Sequence 6, Application US/10213878
; Publication No. US20030073206A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; APPLICANT: Nagarajan, Vasanth
; APPLICANT: Thomas, Stuart
; TITLE OF INVENTION: Use of Xylene Monooxygenase for the Oxidation of Substituted
; TITLE OF INVENTION: Monocyclic Aromatic Compounds
; FILE REFERENCE: CL1662 US NA
; CURRENT APPLICATION NUMBER: US/10/213,878
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 60/311,490
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 12591
; TYPE: DNA
; ORGANISM: Sphingomonas sp.
US-10-213-878-6

Query Match 80.0%; Score 16.8; DB 15; Length 12591;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCCAGG 20
Db 7673 AAGCCTGATGATGCCAGG 7692

RESULT 9

US-10-214-059-6
; Sequence 6, Application US/10214059
; Publication No. US2003007768A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; APPLICANT: Nagarajan, Vasanth
; APPLICANT: Thomas, Stuart
; TITLE OF INVENTION: Use of Xylene Monooxygenase for the Oxidation of Substituted
; TITLE OF INVENTION: Polycyclic Aromatic Compounds
; FILE REFERENCE: CL1663 US NA
; CURRENT APPLICATION NUMBER: US/10/214,059
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 60/311,486
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 12591
; TYPE: DNA
; ORGANISM: Sphingomonas sp.
US-10-214-059-6

Query Match 80.0%; Score 16.8; DB 15; Length 12591;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCCAGG 20
Db 7673 AAGCCTGATGATGCCAGG 7692

RESULT 10

US-09-864-761-11491
; Sequence 11491, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11491
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005739.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
US-09-864-761-11491

Query Match 77.1%; Score 16.2; DB 9; Length 456;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCCAGG 21
Db 15 ATGCTGATTGATTCCAGG 35

RESULT 11

US-09-918-995-283
; Sequence 283, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 283
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(481)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-283

Query Match 77.1%; Score 16.2; DB 11; Length 481;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACGAGG 21

DB 204 AAGCATGATTGAAGACGAGG 224

RESULT 12

US-10-027-632-37978/c
; Sequence 37978, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37978
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-37978

Query Match 77.1%; Score 16.2; DB 13; Length 512;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACGAGG 21

DB 126 AAGCATGTTGATGACGAGG 106

RESULT 13

US-10-027-632-76583/c
; Sequence 76583, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76583
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-76583

Query Match 77.1%; Score 16.2; DB 13; Length 512;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACGAGG 21

DB 126 AAGCATGTTGATGACGAGG 106

RESULT 14

US-10-027-632-37978/c
; Sequence 37978, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37978

; LENGTH: 512
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-37978

Query Match 77.1%; Score 16.2; DB 14; Length 512;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCAAGG 21
|||||
Db 126 AAGCATGGTGGATGACCAAGG 106
|||||

RESULT 15

US-10-027-632-76583/c
; Sequence 76583, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76583
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-76583

Query Match 77.1%; Score 16.2; DB 14; Length 512;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCCTGATTGATGACCAAGG 21
|||||
Db 126 AAGCATGGTGGATGACCAAGG 106
|||||

Search completed: February 20, 2004, 00:50:07
Job time : 197 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:14:23 ; Search time 46 Seconds
(without alignments)
201.501 Million cell updates/sec

Title: US-10-085-108-21_COPY_711_731

Perfect score: 21

Sequence: 1 AAGCTGATTGATGACCGG 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	15.2	72.4	771	4	US-09-252-991A-13016
C 2	15.2	72.4	2325	1	US-08-019-870-2
C 3	15.2	72.4	2325	1	US-08-019-870-4
C 4	15.2	72.4	2325	1	US-08-019-870-7
C 5	15.2	72.4	2325	1	US-08-019-870-10
C 6	15.2	72.4	2325	1	US-08-314-309A-20
C 7	15.2	72.4	2325	1	US-08-633-760-43
C 8	15.2	72.4	2325	1	US-08-633-760-45
C 9	15.2	72.4	2325	1	US-08-633-760-47
C 10	15.2	72.4	2325	1	US-08-633-760-49
C 11	15.2	72.4	2325	1	US-08-633-760-51
C 12	15.2	72.4	2373	1	US-08-019-870-9
C 13	15.2	72.4	2665	3	US-09-040-005-1
C 14	15.2	72.4	2847	1	US-07-747-501A-2
C 15	15.2	72.4	2847	1	US-07-935-312-2
C 16	15.2	72.4	5860	4	US-09-221-017B-820
C 17	15	71.4	541	4	US-09-404-879A-108
C 18	15	71.4	541	4	US-09-338-933-108
C 19	15	71.4	541	4	US-09-215-681-108
C 20	15	71.4	7210	4	US-09-634-238-15
C 21	14.8	70.5	423	1	US-08-470-179-129
C 22	14.8	70.5	423	4	US-09-252-991A-3094
C 23	14.8	70.5	443	1	US-08-592-406-13
C 24	14.8	70.5	591	4	US-09-280-116-138
C 25	14.8	70.5	600	4	US-09-107-532A-2277
C 26	14.8	70.5	654	4	US-09-252-991A-2992
C 27	14.8	70.5	725	3	US-09-197-801-12

C 28	14.8	70.5	725	3	US-09-551-028-12	Sequence 12, Appl
C 29	14.8	70.5	1126	2	US-08-479-733A-27	Sequence 27, Appl
C 30	14.8	70.5	1126	3	US-08-487-427-27	Sequence 27, Appl
C 31	14.8	70.5	1126	3	US-08-479-727A-27	Sequence 27, Appl
C 32	14.8	70.5	1126	3	US-08-482-369A-27	Sequence 27, Appl
C 33	14.8	70.5	1126	5	PCT-US95-07439-27	Sequence 27, Appl
C 34	14.8	70.5	1404	3	US-09-202-101-15	Sequence 15, Appl
C 35	14.8	70.5	1467	4	US-09-367-777-43	Sequence 43, Appl
C 36	14.8	70.5	1467	4	US-09-367-791A-26	Sequence 26, Appl
C 37	14.8	70.5	1500	1	US-08-487-037-4	Sequence 4, Appl
C 38	14.8	70.5	1536	4	US-09-252-991A-3182	Sequence 3182, Ap
C 39	14.8	70.5	1932	4	US-09-640-198D-1	Sequence 1, Appl
C 40	14.8	70.5	2070	4	US-09-252-991A-2687	Sequence 2687, Ap
C 41	14.8	70.5	4080	4	US-09-016-434-1346	Sequence 1346, Ap
C 42	14.8	70.5	5836	4	US-09-453-702B-253	Sequence 253, App
C 43	14.8	70.5	8554	4	US-09-221-017B-523	Sequence 523, App
C 44	14.8	70.5	12848	4	US-09-453-702B-252	Sequence 252, App
C 45	14.8	70.5	13158	2	US-08-687-080-105	Sequence 105, App

ALIGNMENTS

RESULT 1
US-09-252-991A-13016/c
; Sequence 13016, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13016
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13016

Query Match 72.4%; Score 15.2; DB 4; Length 771;
Best Local Similarity 85.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 3;

QY 1 AAGCTTGATTGATGACCGG 20
|||||
Db 208 AAGCTTGTTCTGACCGG 189

RESULT 2
US-08-019-870-2/c
; Sequence 2, Application US/08019870
; Patent No. 5336613
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: YOSHIMASA, SAITO
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: ISHII, YOSHINORI
; TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/019,870
FILING DATE: 19930219
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5336613man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-791-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2322
NAME/KEY: mat_peptide
LOCATION: 4..2322
US-08-019-870-2

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCGTGATTGATGACGAGG 20
Db 1546 AAGCGGATTGGCGACGAGG 1527

RESULT 3
US-08-019-870-4/c
Sequence 4, Application US/08019870
Patent No. 5336613
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: YOSHIMASA, SAITO
APPLICANT: SASAKI, HITOSHI
APPLICANT: ISHII, YOSHINORI
TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/019,870
FILING DATE: 19930219
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5336613man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-791-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2322
NAME/KEY: mat_peptide
LOCATION: 4..2322
US-08-019-870-2
```

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2322
NAME/KEY: mat_peptide
LOCATION: 4..2322
US-08-019-870-4

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCGTGATTGATGACGAGG 20
Db 1546 AAGCGGATTGGCGACGAGG 1527

RESULT 4
US-08-019-870-7/c
Sequence 7, Application US/08019870
Patent No. 5336613
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: YOSHIMASA, SAITO
APPLICANT: SASAKI, HITOSHI
APPLICANT: ISHII, YOSHINORI
TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/019,870
FILING DATE: 19930219
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5336613man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-791-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2322
NAME/KEY: mat_peptide
LOCATION: 4..2322
US-08-019-870-4
```

; LOCATION: 1..2322
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 4..2322
US-08-019-870-7

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCAGG 20
Db 1546 AAGCCGGATTGGCAGCAGG 1527

RESULT 5

US-08-019-870-10/c
; Sequence 10, Application US/08019870
; Patent No. 5336613
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: YOSHIMASA, SAITO
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: ISHII, YOSHINORI
; TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/019,870
; FILING DATE: 19930219
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5336613man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-791-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:

NAME/KEY: CDS
; LOCATION: 1..2322
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 4..2322
US-08-019-870-10

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCAGG 20
Db 1546 AAGCCGGATTGGCAGCAGG 1527

RESULT 6

US-08-314-309A-20/c
; Sequence 20, Application US/08314309A
; Patent No. 5677141
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: FUKAGAWA, MASAO
; APPLICANT: IWAMI, MORITA
; APPLICANT: ARAMORI, ICHIRO
; APPLICANT: KOJO, HITOSHI
; TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEM
; TITLE OF INVENTION: COMPOUND OR SALTS THEREOF
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C. Jefferson Davis Highway, Suite 400
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/314,309A
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631,906
; FILING DATE: 21-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5677141man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-863-0 CONT

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:

NAME/KEY: CDS
; LOCATION: 1..2322
US-08-314-309A-20

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAGCCTGATTGATGACCAGG 20
Db 1546 AAGCCGGATTGGCAGCAGG 1527

RESULT 7

US-08-633-760-43/c
; Sequence 43, Application US/08633760
; Patent No. 5604429
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: FUJIMURA, TAKAO
; APPLICANT: ISHII, YOSHINORI

APPLICANT: NOGUCHI, YUJI
TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,760
FILING DATE: 01-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-929-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2325
FEATURE:
NAME/KEY: mat peptide
LOCATION: 4..2322
US-08-633-760-45

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AGCCTGATTGATGACCAGG 20
Db 1546 AAGCCGGATTGGCAGCAGG 1527

RESULT 8
US-08-633-760-45/C
Sequence 45, Application US/08633760
Patent No. 5804429
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: FUJIMURA, TAKAO
APPLICANT: ISHII, YOSHINORI
TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,760
FILING DATE: 01-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-929-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2325
FEATURE:
NAME/KEY: mat peptide
LOCATION: 4..2322
US-08-633-760-45
Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AGCCTGATTGATGACCAGG 20
Db 1546 AAGCCGGATTGGCAGCAGG 1527
RESULT 9
US-08-633-760-47/C
Sequence 47, Application US/08633760
Patent No. 5804429
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: FUJIMURA, TAKAO
APPLICANT: ISHII, YOSHINORI
TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,760
FILING DATE: 01-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-929-0 PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2325
FEATURE:
NAME/KEY: mat peptide
LOCATION: 4..2322
US-08-633-760-47

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGCCTGATTGACGACCAGG 20
||||| ||||| ||||| |||||
Db 1546 AAGCCGATTGGCGACCAGG 1527

RESULT 10

US-08-633-760-49/c
Sequence 49, Application US/08633760
Patent No. 5804429
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: FUJIMURA, TAKAO
APPLICANT: ISHII, YOSHINORI
APPLICANT: NOGUCHI, YUJI
TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,760
FILING DATE: 01-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-929-0 PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS
LOCATION: 1..2325
FEATURE:
NAME/KEY: mat peptide
LOCATION: 4..2322
US-08-633-760-49

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGCCTGATTGACGACCAGG 20
||||| ||||| ||||| |||||
Db 1546 AAGCCGATTGGCGACCAGG 1527

RESULT 11

US-08-633-760-51/c
Sequence 51, Application US/08633760
Patent No. 5804429
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: FUJIMURA, TAKAO
APPLICANT: ISHII, YOSHINORI
APPLICANT: NOGUCHI, YUJI
TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,760
FILING DATE: 01-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-929-0 PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2325
FEATURE:
NAME/KEY: mat peptide
LOCATION: 4..2322
US-08-633-760-51

Query Match 72.4%; Score 15.2; DB 1; Length 2325;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGCCTGATTGACGACCAGG 20

Db 1546 AAGCGGATTGGCAGCAGG 1527
||||| ||||| |||||

RESULT 12

US-08-019-870-9/c
; Sequence 9, Application US/08019870
; Patent No. 5336613
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: YOSHIMASA, SAITO
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: ISHII, YOSHINOBU
; TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/019,870
; FILING DATE: 19930219
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5336613man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-791-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2373 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-019-870-9

Query Match 72.4%; Score 15.2; DB 1; Length 2373;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCGGATTGGCAGCAGG 20
Db 1594 AAGCGGATTGGCAGCAGG 1575
||||| ||||| |||||

RESULT 13

US-09-040-005-1
; Sequence 1, Application US/09040005
; Patent No. 6057128
; GENERAL INFORMATION:
; APPLICANT: Donaldson, Debra
; APPLICANT: Unger, Michelle
; TITLE OF INVENTION: MU-1 RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA

ZIP: 02140
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,005
; FILING DATE:

CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G15320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-498-8224
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-040-005-1

Query Match 72.4%; Score 15.2; DB 3; Length 2665;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGCCTGATTGATGACGAGG 21
Db 50 AGCCTGATTGATGACGAGG 69
||||| ||||| |||||

RESULT 14

US-07-747-901A-2/c
; Sequence 2, Application US/07747901A
; Patent No. 5192678
; GENERAL INFORMATION:
; APPLICANT: Iwami, Morita
; APPLICANT: Aramori, Ichiro
; APPLICANT: Fukagawa, Masao
; APPLICANT: Isogai, Takao
; APPLICANT: Kojo, Hitoshi
; TITLE OF INVENTION: CEPHALOSPORIN C ACYLASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/747,901A
; FILING DATE: 19910820
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5192678man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-709-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-4500
; TELEFAX: (703) 486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 2847 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 483..2804
US-07-747-901A-2

Query Match 72.4%; Score 15.2; DB 1; Length 2847;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCTGATTGATGACCAGG 20
||||| ||||| ||||| |||||
Db 2028 AAGCCGGATTGGCGACCAGG 2009

RESULT 15

US-07-935-312-2/c
; Sequence 2, Application US/07935312
; Patent No. 5320948
; GENERAL INFORMATION:
; APPLICANT: Iwami, Morita
; APPLICANT: Aramori, Ichiro
; APPLICANT: Fukagawa, Masao
; APPLICANT: Isogai, Takao
; APPLICANT: Kojo, Hitoshi
; TITLE OF INVENTION: CEPHALOSPORIN C ACYLASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,312
; FILING DATE: 19920826
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5320948man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-769-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2847 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 483..2804
US-07-935-312-2

Query Match 72.4%; Score 15.2; DB 1; Length 2847;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGCTGATTGATGACCAGG 20
||||| ||||| ||||| |||||

Db 2028 AAGCCGGATTGGCGACCAGG 2009

Search completed: February 19, 2004, 23:16:14
Job time : 47 secs

